内 容 提 要

本书是研究生随机过程教材. 全书共 4 章, 以公理概率论为入口, 重点讲授鞅与 Markov 过程, 分别介绍了条件期望、无穷维空间的测度构造、Markov 链、Poisson 测度与 Poisson 过程、Brown 运动、鞅与连续鞅的随机积分、Itô 公式、Girsanov 公式、随机微分方程, 还介绍了右 Markov 过程、Feller 过程与 Lévy 过程、Brown 运动的位势理论、游离理论, 和 Markov 过程的 Killing 变换与时间变换等. 本书还配备了一定数量难易不等的习题, 以利读者加深理解, 启发思考.

本书可作为基础数学、应用数学、计算数学、运筹学与控制论、概率论与数理统计等数学类各专业方向的研究生学位课教材, 也可供理工类和金融类相关专业的研究生以及自然科学工作者、工程技术人员参考使用.

编辑出版说明

21世纪, 随着科学技术的突飞猛进和知识经济的迅速发展, 世界将发生深刻变化, 国际间的竞争日趋激烈, 高层次人才的教育正面临空前的发展机遇与巨大挑战.

研究生教育是教育结构中最高层次的教育, 肩负着为国家现代化建设培养高素质、高层次创造性人才的重任, 是我国增强综合国力、增强国际竞争力的重要支撑, 为了提高研究生的培养质量和研究生教学的整体水平, 必须加强研究生的教材建设, 更新教学内容, 把创新能力和创新精神的培养放在突出位置上, 必须建立适应新的教学和科研要求的有复旦特色的研究生教学用书,"21世纪复旦大学研究生教学用书"正是为适应这一新形势而编辑出版的.

"21世纪复旦大学研究生教学用书"分文科、理科和医科三大类, 主要出版硕士研究生学位基础课和学位专业课的教材, 同时酌情出版一些使用面广、质量较高的选修课及博士研究生学位基础课教材, 这些教材除可作为相关学科的研究生教学用书外, 还可供有关学者和人员参考.

收入"21世纪复旦大学研究生教学用书"的教材, 大多是作者在编写成讲义后, 经过多年的教学实践、反复修改后才定稿的. 这些作者大多治学严谨, 教学实践经验丰富, 教学效果也比较显著. 由于我们对编辑工作尚缺乏经验, 不足之处, 敬请读者指正, 以便我们在将来再版时加以更正和提高.

复旦大学研究生院

前　言

概率论是研究机会的学科, 有非常强的直观背景, 它起源于人类对赌博中机会的兴趣, 对它的一些问题的研究可追溯至几个世纪之前, 著名数学家 P. Fermat, B. Pascal, J. Bernoulli, P.S. Laplace 等都对概率论的发展作出过巨大贡献, 而为这一学科建立坚实的数学基础是 20 世纪 30 年代, 俄国著名数学家 A.N.Kolmogorov 将概率论的大厦建立在测度论的基石上. 但我们不能忘记概率与测度不同的一面, 它有着深刻的且是严格的测度论公理体系所无法体现的直观背景.

此教材重点讲述 Markov 过程与鞅论, 考虑到大多数学生在本科时缺乏测度论的训练, 我们在第一章简要介绍测度论与概率论的基本概念与重要结果, 如 Caratheodory 扩张定理、Radon-Nikodym 定理、随机变量及其分布、条件数学期望等, 要注意的是没有对这些概念的真正理解, 是不可能真正理解现代随机过程理论的. 在第二章中, 我们将介绍 Kolmogorov 的相容性定理, 以及一些常见的过程, 如平稳过程、Markov 链、Markov 过程、独立增量过程、Poisson 过程、Brown 运动等. 在第三章中, 我们将给出鞅的定义, 讨论鞅的基本性质、鞅不等式、鞅与其他随机过程的关系, 以及鞅的正则化. 另外我们还将证明连续鞅有有限二次变差并定义随机积分, 最后介绍重要的 Itô 公式及其应用.

该教材在浙江大学和复旦大学作为数学系研究生随机过程基础课的教材已经使用多年, 此次出版前增加并修订了一些内容. 考虑到 Markov 过程的重要性, 还增加了第四章: Markov 过程基础, 讲述 Markov 过程、强 Markov 性、概率位势理论、Feller 过程和 Lévy 过程等, 这部分内容对于数学系研究生基础课可能过于专门化, 适合作为概率专业研究生的专业课内容讲授.

这里要感谢浙江大学的陈叔平教授, 他首先提议和鼓励我们为浙大数学系研究生开设这门课程并提供方便, 感谢赵敏智、方兴、张慧增、何萍博士, 他们多次阅读此教材并为教材的修改提出了许多的重要意见, 感谢周梦、吴小伟同学, 他们在阅读过程中也指出并改正了一些错误. 我们还要特别感谢汪嘉冈教授、马志明教授和李贤平教授, 他们在百忙中仔细地阅读了全书并提出了许多重要的修改意见. 最后还要感谢复旦大学出版社的范仁梅女士为本书顺利出版提供的帮助. 教材虽经不断的修改, 但错误依然难免, 如果读者发现其中的错误或有建议, 请直接和作者联系, 非常感谢.

应坚刚　jgying@fudan.edu.cn

金蒙伟　jmw@zju.edu.cn

目　录

第一章 概率论基础

在这一章中, 我们将考察概率论的基本概念, 它们也是随机分析理论的基础. 我们将简略而又系统地介绍测度论与概率论的重要概念和定理, 为了让此书尽量自我包含, 我们将简要地给予证明. 需要强调的是, 虽然这里我们在测度论基础上建立概率的概念, 但是概率的直观思想是远非测度论所能体现的, 所以读者不能省略对于初等概率论的系统学习和理解.

§1.1 可测结构与测度构造

在本书中, 集合 $\mathbf{R}$ 表示实数集, $\mathbf{Q}$ 表示有理数集, $\mathbf{Z}$ 表示整数集, $\mathbf{N}$ 表示自然数集, 下标 $+$ 表示非负元素全体, 如 $\mathbf{R}_+$ 表示非负实数集, 其他类似.

用 Ω 表示一个任意给定的非空集合, 2^Ω 表示 Ω 的子集全体组成的集合, 称为幂集, Ω 的一个子集类是指 2^Ω 的一个子集. 我们说一个子集类对集合的某种运算封闭, 是指此子集类中的集合经过此种运算后得到的集合还在此子集类内. 常用的集合运算如下:

(1) 补集: $A \mapsto A^c = \Omega \setminus A$;

(2) 有限并: $(A, B) \mapsto A \cup B$; 可列并: $(\{A_n : n \in \mathbf{N}\}) \mapsto \bigcup\limits_n A_n$; 任意并: $(\{A_\lambda : \lambda \in \Lambda\}) \mapsto \bigcup\limits_\lambda A_\lambda$; 当上面运算中所涉及的集合不相交时, 分别称为不交有限并, 不交可列并, 不交任意并;

(3) 有限交: $(A, B) \mapsto A \cap B$; 可列交: $(\{A_n : n \in \mathbf{N}\}) \mapsto \bigcap\limits_n A_n$; 任意交: $(\{A_\lambda : \lambda \in \Lambda\}) \mapsto \bigcap\limits_\lambda A_\lambda$;

(4) 差: $(A, B) \mapsto A \setminus B = A \cap B^c$; 当 $A \supset B$ 时, 称为包含差;

(5) 递增列极限: $(\{A_n : n \in \mathbf{N}\}) \mapsto \bigcup\limits_n A_n$, 其中集列 $\{A_n\}$ 递增;

(6) 递减列极限: $(\{A_n : n \in \mathbf{N}\}) \mapsto \bigcap\limits_n A_n$, 其中集列 $\{A_n\}$ 递减.

我们期望读者已经熟悉关于集合运算的规则, 这里不一一列举.

定义 1.1.1 Ω 的一个非空子集类 $\mathscr{F}$ 称为是 Ω 上的 σ- 代数 (或 σ- 域), 如果它对于补集运算和可列并运算封闭.

容易验证, σ- 代数一定包含有 $\varnothing, \Omega$ 为元素且对于有限交, 有限并及可列交等运算都是封闭的. 集合上的一个 σ- 代数通常看作为集合上的可测结构, Ω及其上的一个 σ- 代数 $\mathscr{F}$ 组成的偶 $(\Omega, \mathscr{F})$ 称为是一个可测空间. 显然子集类 2^Ω 与 $\{\varnothing, \Omega\}$ 是 Ω上的 σ- 代数, 它们是 Ω 上的平凡 σ- 代数.

例 1.1.1 设 $\mathscr{F}$ 是实数集 $\mathbf{R}$ 的至多可列子集或补集是至多可列子集全体, 那么 $\mathscr{F}$ 是 $\mathbf{R}$ 上的一个 σ- 代数. ▮

由定义不难验证 Ω 上任意多个 σ- 代数的交也是一个 σ- 代数, 设 $\mathscr{A}$ 是 Ω 上一个子集类, 用 $C(\mathscr{A})$ 表示 Ω 上包含 $\mathscr{A}$ 为子集的 σ- 代数全体, 因为 $2^\Omega \in C(\mathscr{A})$, 故 $C(\mathscr{A})$ 是非空的, 记

$$\sigma(\mathscr{A}) := \bigcap_{\mathscr{F} \in C(\mathscr{A})} \mathscr{F},$$

(在本书中, 记号 := 读作 被定义为.) 则 $\sigma(\mathscr{A})$ 是 Ω 上的 σ- 代数, 它是由下列条件所唯一确定的 σ- 代数 $\mathscr{F}$:

(1) $\mathscr{F} \supset \mathscr{A}$;

(2) 若 $\mathscr{F}'$ 是 σ- 代数且 $\mathscr{F}' \supset \mathscr{A}$, 则 $\mathscr{F}' \supset \mathscr{F}$.

因此称 $\sigma(\mathscr{A})$ 是包含 $\mathscr{A}$ 的最小 σ- 代数或由 $\mathscr{A}$ 生成的 σ- 代数. 这是生成大多数重要的 σ- 代数的常用方法.

例 1.1.2 如果 Ω 是一个拓扑空间, 则其所有开集组成的集类生成的 σ- 代数称为是 Ω 上的 Borel 代数, 记为 $\mathscr{B}(\Omega)$, 因为开集的补集是闭集, 故它也是全体闭集生成的 σ- 代数. 对于 Euclid 空间, 我们记 $\mathscr{B}(\mathbf{R}^n)$ 或 $\mathscr{B}^n$ 是 n- 维 Euclid 空间 $\mathbf{R}^n$ 上的 Borel 代数. 一个 σ- 代数 $\mathscr{A}$ 是可列生成的, 如果存在子集列 $\{A_n\}$ 使得 $\mathscr{A} = \sigma(\{A_n\})$. 那么 $\mathscr{B}(\mathbf{R}^n)$ 是可列生成的. ▮

定义 1.1.2 设 $(\Omega, \mathscr{F})$ 是一个可测空间, 称 $\mathscr{F}$ 上的一个非负实值广义(可取无穷值的)集函数 μ 为 $(\Omega, \mathscr{F})$ 上的测度, 如果

(1) $\mu(\varnothing) = 0$;

(2) 若 $\{A_n\}$ 是 $\mathscr{F}$ 中的一个互不相交的集列, 则 $\mu(\bigcup_n A_n) = \sum_n \mu(A_n)$. 这个性质称为测度的可列可加性.

这时, 称 $(\Omega, \mathscr{F}, \mu)$ 是测度空间. 如果 $A \in \mathscr{F}, \mu(A) = 0, B \subset A$ 蕴含着 $B \in \mathscr{F}$, 则称 $(\Omega, \mathscr{F}, \mu)$ 是完备测度空间. 当 $\mu(\Omega) < \infty$ 时, 称 μ 是有限测度; 当 $\mu(\Omega) = 1$ 时, 称 μ 为概率测度; 当存在集列 $\{A_n\} \subset \mathscr{F}$ 满足 $\bigcup_n A_n = \Omega$ 与 $\mu(A_n) < \infty$ 时, 称 μ 是 σ- 有限测度. 另外如果 μ 是拓扑空间 Ω 及其 Borel 集上的测度, 如果 μ 在任何紧集上有限, 称 μ 为Radon 测度.

可测空间 $(\Omega, \mathscr{F})$ 的测度有一个偏序, 称测度 $\nu \leqslant \mu$, 如果对任何 $A \in \mathscr{F}$, $\nu(A) \leqslant \mu(A)$. 另外, 任何测度空间在下面的意义下可以完备化. 设 $(\Omega, \mathscr{F}, \mu)$ 是测度空间. 记

$$\mathscr{N} := \{N \subset \Omega: \text{存在 } N' \in \mathscr{F} \text{ 使得 } N \subset N', \mu(N') = 0\},$$

$$\mathscr{F}^\mu := \sigma(\mathscr{F} \cup \mathscr{N}).$$

$\mathscr{N}$ 中的集合通常称为 μ- 零测集, 那么 $\mathscr{F}^\mu = \{A \cup N : A \in \mathscr{F}, N \in \mathscr{N}\}$, 这样 μ 自动地延拓到 $\mathscr{F}^\mu$ 上: $\mu(A \cup N) := \mu(A)$. 读者需要验证定义无歧义, 则 $(\Omega, \mathscr{F}^\mu, \mu)$ 是一个完备测度空间, 称为是原测度空间的完备化. 因此如有必要, 我们总可以假设测度空间是完备的.

下面有关测度的性质可由定义直接推得.

(1) (有限可加性) 若 $A_1, \cdots, A_n \in \mathscr{F}$ 且互不相交, 那么

$$\mu\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mu(A_i);$$

(2) (单调性) 若 $A, B \in \mathscr{F}$ 且 $A \subset B$, 则 $\mu(A) \leqslant \mu(B)$;

(3) (次可列可加性) 若 $\{A_n\}$ 是 $\mathscr{F}$ 中的一个集列, 那么

$$\mu\left(\bigcup_n A_n\right) \leqslant \sum_n \mu(A_n);$$

(4) (下连续性) 设 $\{A_n\}$ 是 $\mathscr{F}$ 中单调上升的集列, 则

$$\mu(\lim A_n) = \lim \mu(A_n);$$

(5) (上连续性) 设 $\{A_n\}$ 是 $\mathscr{F}$ 中单调下降的集列且存在 k 使得 $\mu(A_k) < \infty$, 则

$$\mu(\lim A_n) = \lim \mu(A_n).$$

特别地, 在 $\varnothing$ 处上连续, 即若 $\{A_n\}$ 是 $\mathscr{F}$ 中单调下降交为 $\varnothing$ 的集列且存在 k, 使得 $\mu(A_k) < \infty$, 则 $\lim \mu(A_n) = 0$.

注意有限可加性与次可列可加性两者结合等价于可列可加性.

现在, 我们来看几个简单的常用测度. 恒等于零的测度称为是零测度. 在空集上等于零, 而在非空可测集上等于 $+\infty$ 的集函数也是一个测度.

例 1.1.3 设 Ω 是非空集, 对任意 $A \subset \Omega$, 用 $\#(A)$ 表示集合 A 中元素的个数, 显然 $\#$ 是 $(\Omega, 2^\Omega)$ 上的测度, 且当 Ω 是有限集时, 它是有限测度, 当 Ω 是可列集时, 它是 σ- 有限测度, 而当 Ω 不可列时, 它不是 σ- 有限的. 测度 $\#$ 通常称为 Ω 上的计数测度. ▮

例 1.1.4 设 $(\Omega, \mathscr{F})$ 是可测空间, 对任意 $\omega \in \Omega$, 定义

$$\epsilon_\omega(A) := 1_A(\omega),\ A \in \mathscr{F}.$$

同样, 显然 ϵ_ω 是测度, 通常被称为 ω 处的单点测度或 Dirac 测度.

测度的定义并非总是如此简单, 实际上, 当集合上的可测结构较为复杂时, 像上面例中那样直接对每个可测集定义而成为测度是不可能的. 因此我们通常是在一个相对简单的集类上直接地定义一个(预)测度, 然后用某种方法将其延拓至该集类生成的 σ- 代数上, 这正是下面将介绍的著名的 Carathéodory 测度扩张定理的主要思想.

定义 1.1.3 Ω 的非空子集类 $\mathscr{F}_0$ 称为是一个环, 如果 $\mathscr{F}_0$ 对集合的差与有限并两种运算封闭.

显然环含有空集且对于有限交运算也封闭.

例 1.1.5 任何非空集合的有限子集全体是一个环. 定义 $\mathbf{R}$ 上的子集类如下

$$\mathscr{B}_0 := \left\{ \bigcup_{i=1}^n (a_i, b_i] :\ n \geqslant 0, a_i \leqslant b_i, 1 \leqslant i \leqslant n \right\},$$

则 $\mathscr{B}_0$ 是 $\mathbf{R}$ 上的一个环, 它生成的 σ- 代数也是 Borel 代数. ▮

可以看出, 环的结构通常比 σ- 代数的结构简单得多, 故在环上定义一个测度也比在 σ- 代数上定义测度要简单得多.

定义 1.1.4　设 $\mathscr{F}_0$ 是非空集合 Ω 上的一个环, $\mathscr{F}_0$ 上的非负广义实值集函数 μ 称为是有限可加集函数, 如果下列条件满足:

(1) $\mu(\varnothing)=0$;

(2) (有限可加) 若 $A_1,\cdots,A_n\in\mathscr{F}_0$ 互不相交, 则

$$\mu\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}\mu(A_i).$$

进一步地, 如果将(2)用(2′)代之,

(2′) (可列可加) 若集列 $\{A_n\}\subset\mathscr{F}_0$ 不相交且 $\bigcup_n A_n\in\mathscr{F}_0$, 则

$$\mu\left(\bigcup_{n}A_n\right)=\sum_{n}\mu(A_n),$$

则称 μ 是预测度. 记 $\mathscr{F}:=\sigma(\mathscr{F}_0)$, 如果 $\overline{\mu}$ 是 $(\Omega,\mathscr{F})$ 上的测度且其在 $\mathscr{F}_0$ 上与 μ 一致, 称 $\overline{\mu}$ 是 μ 上的扩张.

显然, 只要在集合运算封闭的前提下, 测度的性质可以推广到环上的预测度上, 读者不妨自行验证. 在概率论中, 我们更常使用的是代数的结构, Ω 上的一个环 $\mathscr{F}_0$ 称为是代数, 如果 $\Omega\in\mathscr{F}_0$.

设 $\mathscr{F}_0$ 是 Ω 上的环, μ 是 $\mathscr{F}_0$ 上给定的预测度, 对 $A\subset\Omega$, 定义

$$\mu^*(A):=\inf\left\{\sum_{n}\mu(A_n):\ \{A_n\}\subset\mathscr{F}_0,\ \bigcup_{n}A_n\supset A\right\}$$

(约定 $\inf\varnothing=+\infty$), 自然 μ^* 是 2^{Ω} 上的非负广义实值集函数, 称其为(由 μ 诱导的)外测度, 它有下列性质:

(1) 若 $A\in\mathscr{F}_0$, 则 $\mu^*(A)=\mu(A)$;

(2) (单调性) 若 $A\subset B$, 则 $\mu^*(A)\leqslant\mu^*(B)$;

(3) (次可列可加性) 若 $\{A_n\}$ 是 Ω 的一个子集列, 那么

$$\mu^*\left(\bigcup_{n}A_n\right)\leqslant\sum_{n}\mu^*(A_n).$$

其中(2), (3) 可由定义直接验证, 而 (1) 的证明要用到 μ 的次可列可加性.

一般地, 2^{Ω} 上的一个空集上为零且满足 (2), (3) 的非负集函数常称为是 Ω 上

的外测度. Ω 的任意子集 A 称为是 μ^*- 可测的, 如果对任何 $E\subset\Omega$, 有

$$\mu^*(E)=\mu^*(A\cap E)+\mu^*(A^c\cap E).$$

由外测度的性质, 上式等价于

$$\mu^*(E)\geqslant\mu^*(A\cap E)+\mu^*(A^c\cap E).$$

记 $\mathscr{M}$ 为 Ω 的 μ^*- 可测子集全体.

定理 1.1.1 (Carathéodory) 定义如上, 则有下列性质:

(1) $(\Omega,\mathscr{M},\mu^*)$ 是测度空间;

(2) $\mathscr{M}\supset\mathscr{F}_0$;

(3) $\mathscr{M}$ 包含所有 μ^*- 零测集. 因此, $(\Omega,\mathscr{M},\mu^*)$ 是完备的.

证明 (1) 我们须证 $\mathscr{M}$ 是 Ω 上 σ- 代数且 μ^* 是 $\mathscr{M}$ 上的测度. 显然 $\varnothing,\Omega\in\mathscr{M}$ 且 $\mathscr{M}$ 对补集运算封闭, 故只须验证 $\mathscr{M}$ 对可列并运算封闭. 事实上, $\mathscr{M}$ 对有限并运算封闭, 因为若 $A,B\in\mathscr{M}$, 则对 $E\subset\Omega$,

$$\begin{aligned}\mu^*(E)&=\mu^*(A\cap E)+\mu^*(A^c\cap E)\\&=\mu^*(A\cap E)+[\mu^*(B\cap A^c\cap E)+\mu^*(B^c\cap A^c\cap E)]\\&=[\mu^*(A\cap(A\cup B)\cap E)+\mu^*((A\cup B)\cap A^c\cap E)]\\&\quad+\mu^*((B\cup A)^c\cap E)\\&=\mu^*((A\cup B)\cap E)+\mu^*((B\cup A)^c\cap E),\end{aligned}$$

推出 $A\cup B\in\mathscr{M}$. 那么 $\mathscr{M}$ 对有限交运算也封闭. 故我们仅须验证 $\mathscr{M}$ 对不相交集列的可列并运算封闭. 设 $\{A_n\}$ 是 $\mathscr{M}$ 中不相交集列, 令

$$B_n:=\bigcup_{i=1}^n A_i,\ A:=\bigcup_{n=1}^\infty A_n,$$

由外测度的单调性得

$$\begin{aligned}\mu^*(E)&=\mu^*(E\cap B_n)+\mu^*(E\cap B_n^c)\\&\geqslant\mu^*(E\cap B_n)+\mu^*(E\cap A^c)\\&=\mu^*(E\cap B_n\cap B_{n-1})+\mu^*(E\cap B_n\cap B_{n-1}^c)+\mu^*(E\cap A^c)\\&=\mu^*(E\cap B_{n-1})+\mu^*(E\cap A_n)+\mu^*(E\cap A^c)\end{aligned}$$

$$= \cdots = \sum_{i=1}^{n} \mu^*(E \cap A_i) + \mu^*(E \cap A^c).$$

由 n 的任意性与 μ^* 的次可列可加性推出

$$\mu^*(E) \geqslant \sum_{i=1}^{\infty} \mu^*(E \cap A_i) + \mu^*(E \cap A^c)$$
$$\geqslant \mu^*(E \cap A) + \mu^*(E \cap A^c),$$

因此 $A \in \mathscr{M}$. 不仅如此, 从上面的证明过程中可以看出对任何 $E \subset \Omega$,

$$\mu^*\left(E \cap (\bigcup_{i=1}^{n} A_i)\right) = \sum_{i=1}^{n} \mu^*(E \cap A_i),$$

故 μ^* 有有限可加性

$$\mu^*\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu^*(A_i).$$

结合 μ^* 的次可列可加性, 推出 μ^* 在 $\mathscr{M}$ 上有可列可加性. 因此 μ^* 是 $(\Omega, \mathscr{M})$ 上的测度, 显然, 证明对于一般的外测度都对.

(2) 设 $A \in \mathscr{F}_0$, 对任何 $E \subset \Omega$, 不妨设 $\mu^*(E) < \infty$, 则对任何 $\epsilon > 0$, 存在子集列 $\{A_n\} \subset \mathscr{F}_0$ 满足 $\bigcup\limits_n A_n \supset E$ 且 $\sum\limits_n \mu(A_n) < \mu^*(E) + \epsilon$, 因而

$$\begin{aligned}
&\mu^*(E \cap A) + \mu^*(E \cap A^c) \\
\leqslant\ & \mu^*((\bigcup_n A_n) \cap A) + \mu^*((\bigcup_n A_n) \cap A^c) \\
=\ & \mu^*(\bigcup_n (A_n \cap A)) + \mu^*(\bigcup_n (A_n \cap A^c)) \\
\leqslant\ & \sum_n [\mu^*(A_n \cap A) + \mu^*(A_n \cap A^c)] \\
\leqslant\ & \sum_n [\mu(A_n \cap A) + \mu(A_n \cap A^c)] \\
=\ & \sum_n \mu(A_n) < \mu^*(E) + \epsilon,
\end{aligned}$$

其中最后的等式来自 μ 的有限可加性. 由此推出 $\mu^*(E\cap A)+\mu^*(E\cap A^c)\leqslant\mu^*(E)$, 即 $A\in\mathscr{M}$ 或 $\mathscr{F}_0\subset\mathscr{M}$.

(3) 设 A 是 μ^*- 零测集, 即 $\mu^*(A)=0$, 则对任何 $E\subset\Omega$, 有 $\mu^*(E)\leqslant\mu^*(E\cap A^c)$, 因此 $\mu^*(E)=\mu^*(E\cap A^c)=\mu^*(E\cap A)+\mu^*(E\cap A^c)$, 故 $A\in\mathscr{M}$. □

由定理推出 $\mathscr{M}\supset\mathscr{F}$, 则 μ^* 限制在 $\mathscr{F}$ 上是 μ 的一个扩张, 称它为 μ 的 Carathéodory 扩张, 在不至于引起误解时, 仍记为 μ. 注意, σ- 代数 $\mathscr{M}$ 与 $\mathscr{F}_0$ 和 μ 都有关, 而 $\mathscr{F}$ 仅与 $\mathscr{F}_0$ 有关. $(\Omega,\mathscr{M},\mu^*)$ 是完备测度空间, 而 $(\Omega,\mathscr{F},\mu)$ 一般不是完备的. 环上预测度的 Carathéodory 扩张不一定是唯一的扩张.

例 1.1.6 取例 1.1.5 中 $\mathbf{R}$ 上的环 $\mathscr{B}_0$, 对 $A\in\mathscr{B}_0$, 定义

$$\mu(A):=\begin{cases}0, & A=\varnothing,\\ +\infty, & A\neq\varnothing,\end{cases}$$

则 μ 的 Carathéodory 扩张是 $(\mathbf{R},\mathscr{B})$ 上的奇异测度, 而显然计数测度也是 μ 的一个扩张, 它不同于 Carathéodory 扩张.

下面我们将证明预测度的 σ- 有限性能保证 Carathéodory 扩张是唯一扩张. 称环 $\mathscr{F}_0$ 上预测度 μ 是 σ- 有限的, 如果存在集列 $\{\Omega_n\}\subset\mathscr{F}_0$ 满足 $\bigcup\limits_n\Omega_n=\Omega$ 且 $\mu(\Omega_n)<\infty$. 设 $(\Omega,\mathscr{F},\mu)$ 是测度空间, 对 $\Omega_0\in\mathscr{F}$, 定义 $\mathscr{F}\cap\Omega_0:=\{\Omega_0\cap A: A\in\mathscr{F}\}$, 则 $\mathscr{F}\cap\Omega_0$ 是 Ω_0 上的 σ- 代数, 记 μ_{Ω_0} 是 μ 在 $\mathscr{F}\cap\Omega_0$ 上的限制, 那么 $(\Omega_0,\mathscr{F}\cap\Omega_0,\mu_{\Omega_0})$ 也是一个测度空间.

我们还需要介绍 Dynkin 的一个思想, 它给出证明一个子集类是 σ- 代数的一个方法, 在下面定理的证明中会使用它. 称一个子集类是π- 类, 如果它对有限交封闭. 而称一个子集类是 Dynkin 系或 λ- 类, 如果它包含有 $\varnothing,\Omega$ 且对于补集运算与不交可列并运算封闭. 显然, 环是 π- 类, σ- 代数是 Dynkin 系, 反之不对. 容易看出任意多个 Dynkin 系的交仍是 Dynkin 系, 因此对 Ω 的任何子集类 $\mathscr{A}$, 唯一存在一个包含 $\mathscr{A}$ 的最小 Dynkin 系, 记为 $\delta(\mathscr{A})$, 也类似地称为由 $\mathscr{A}$ 生成的 Dynkin 系. 下面是 Dynkin 定理.

引理 1.1.1 (Dynkin) 设 $\mathscr{F}_0$ 是一个 π- 类, 则 $\delta(\mathscr{F}_0)$ 是一个 σ- 代数, 因此 $\sigma(\mathscr{F}_0)=\delta(\mathscr{F}_0)$.

证明　由定义, 仅须验证 $\delta(\mathscr{F}_0)$ 对有限交运算封闭. 任取 $A \in \delta(\mathscr{F}_0)$, 定义

$$\kappa[A] := \{B \in \delta(\mathscr{F}_0):\ A \cap B \in \delta(\mathscr{F}_0)\}.$$

先验证 $\kappa[A]$ 是一个 Dynkin 系. 事实上, 只需证明 $\kappa[A]$:

(1) 对补集运算封闭;

(2) 对不相交集列的可列并运算封闭.

对 (1), 取 $B \in \kappa[A]$, 则 $A, A^c, A \cap B \in \delta(\mathscr{F}_0)$, 因此 $A \cap B^c = [A^c \cup (A \cap B)]^c \in \delta(\mathscr{F}_0)$, 因此 $B^c \in \kappa[A]$.

为证明 (2), 取 $\{B_n\} \subset \kappa[A]$ 是不交集列, 则显然 $\{A \cap B_n\}$ 是 $\delta(\mathscr{F}_0)$ 中不交集列, 因此 $A \cap (\bigcup_n B_n) \in \delta(\mathscr{F}_0)$, 推出 $\bigcup_n B_n \in \kappa[A]$.

因 $\mathscr{F}_0$ 是 π- 类, 故 $A \in \mathscr{F}_0$ 蕴含着 $\kappa[A] \supset \mathscr{F}_0$ 即 $\kappa[A] \supset \delta(\mathscr{F}_0)$. 这意味着当 $A \in \delta(\mathscr{F}_0)$ 时, $\kappa[A] \supset \mathscr{F}_0$. 因此 $\kappa[A] \supset \delta(\mathscr{F}_0)$, 即 $\delta(\mathscr{F}_0)$ 中元素对有限交运算封闭. □

定理 1.1.2　如果 μ 是环 $\mathscr{F}_0$ 上的 σ- 有限的预测度, 则 μ 在最小 σ- 代数 $\mathscr{F}$ 上的扩张是唯一的.

证明　任取 μ 的一个扩张 $\overline{\mu}$, 取 $\Omega_0 \in \mathscr{F}_0$ 且 $\mu(\Omega_0) < \infty$. 令

$$\mathscr{A}_0 := \{A \in \mathscr{F} \cap \Omega_0:\ \overline{\mu}(A) = \mu(A)\}.$$

因 $\overline{\mu}$ 与 μ 在 $\mathscr{F} \cap \Omega_0$ 上是有限测度, 故容易验证: (i) $\mathscr{A}_0 \supset \mathscr{F}_0 \cap \Omega_0$; (ii) $\mathscr{A}_0$ 是 Ω_0 上的 Dynkin 系. 因 $\mathscr{F}_0 \cap \Omega_0$ 是 Ω_0 上的环, 故 $\mathscr{A}_0 \supset \delta(\mathscr{F}_0 \cap \Omega_0) = \sigma(\mathscr{F}_0 \cap \Omega_0) = \mathscr{F} \cap \Omega_0$. 因此 $\overline{\mu}$ 与 μ 在 $\mathscr{F} \cap \Omega_0$ 上一致.

由 σ- 有限性, 可取递增集列 $\{\Omega_n\} \subset \mathscr{F}_0$ 满足 $\bigcup_n \Omega_n = \Omega$ 且 $\mu(\Omega_n) < \infty$. 那么对任何 $A \in \mathscr{F}$, $\overline{\mu}(A) = \lim_n \overline{\mu}(A \cap \Omega_n) = \lim \mu(A \cap \Omega_n) = \mu(A)$. □

上面的几个定理说明一个测度空间可以在环上构造得到, 但实际上, 这通常可以在更简单的半环上进行(参见习题). 在例 1.1.5 中, 定义 $m((a,b]) = b - a$, 则仿照下面定理 1.3.1 的证明, m 是 $\mathscr{B}_0$ 上的一个预测度, 所诱导的 m^*- 可测集称为 Lebesgue 可测集, 记此集合为 $\mathscr{L}$, 诱导的测度称为 Lebesgue 测度. 注意:

(1) 不是所有 $\mathbf{R}$ 的子集都是 Lebesgue 可测的, 或者说一个平移不变的可列可加集函数是不可能在全体子集上定义的;

(2) $\mathscr{L}$ 严格地包含 Borel 代数 $\mathscr{B}$, 实际上, $\mathscr{L}$ 与 $\mathbf{R}$ 的幂集一一对应而 $\mathscr{B}$ 与 $\mathbf{R}$ 一一对应.

设 Ω, Ω' 是两个非空集合, ξ 是 Ω 到 Ω' 的一个映射. 对 $A' \subset \Omega'$, 定义

$$\xi^{-1}(A') := \{\omega \in \Omega : \xi(\omega) \in A'\},$$

则容易验证下列性质:

(1) $\xi^{-1}(\Omega') = \Omega$, $\xi^{-1}(\varnothing) = \varnothing$;

(2) $\xi^{-1}[(A')^c] = [\xi^{-1}(A')]^c$;

(3) 对任何子集列 $\{A'_i\}$, $\xi^{-1}\left(\bigcup_i A'_i\right) = \bigcup_i \xi^{-1}(A'_i)$.

设 $\mathscr{A}'$ 是 Ω' 的一个子集类, 令

$$\xi^{-1}(\mathscr{A}') := \{\xi^{-1}(A') : A' \in \mathscr{A}'\},$$

则由上述性质, 如果 $\mathscr{A}'$ 是 Ω' 上 σ- 代数, $\xi^{-1}(\mathscr{A}')$ 是 Ω 上 σ- 代数. 可测空间 $(\Omega, \mathscr{F})$ 到可测空间 $(\Omega', \mathscr{F}')$ 的映射 ξ 称为是可测的(或明确地, $\mathscr{F}/\mathscr{F}'$- 可测的), 如果 $\xi^{-1}(\mathscr{F}') \subset \mathscr{F}$. 设 $\{\xi_\lambda : \lambda \in \Lambda\}$ 是 Ω 到 Ω' 的映射族, $\mathscr{F}'$ 是 Ω' 上的 σ- 代数, 那么 Ω 上存在唯一一个使得映射 $\{\xi_\lambda\}$ 都可测的最小 σ- 代数 $\mathscr{F}$, 即 (i) 每个 ξ_λ 是可测映射; (ii) 如果 Ω 上另外一个 σ- 代数 $\mathscr{F}_1$ 使得每个 ξ_λ 都可测, 那么 $\mathscr{F} \subset \mathscr{F}_1$. 事实上, 不难验证 $\mathscr{F}$ 是由 $\bigcup\limits_{\lambda \in \Lambda} \xi_\lambda^{-1}(\mathscr{F}')$ 所生成的 σ- 代数, 记为 $\sigma(\{\xi_\lambda : \lambda \in \Lambda\})$.

设 μ 是 $(\Omega, \mathscr{F})$ 上的一个测度, 则可测映射 ξ 把 μ 映为 $(\Omega', \mathscr{F}')$ 上的测度 $\mu \circ \xi^{-1}$ (或记为 $\xi(\mu)$):

$$\xi(\mu)(A') := \mu(\xi^{-1}(A')),\ A' \in \mathscr{F}'.$$

称为 μ 在 ξ 下的像测度.

下面介绍 Hahn 分解和 Jordan 分解. 设 $(\Omega, \mathscr{F})$ 是可测空间, 称 $\mathscr{F}$ 上广义实值集函数 μ 为一个符号测度, 如果: (i) $\mu(\varnothing)=0$; (ii) μ 满足可列可加性. (其中, 可加性蕴含着或者对任何 $A \in \mathscr{F}$, $\mu(A) < +\infty$ 或者对任何 $A \in \mathscr{F}$, $\mu(A) > -\infty$.) 称符号测度是有限的, 如果对任何 $A \in \mathscr{F}$, $|\mu(A)| < +\infty$. 集 $A \in \mathscr{F}$ 是正集(负集), 如果其任何可测子集测度非负(非正). 以下习题中, 设 μ 是 $(\Omega, \mathscr{F})$ 上符号测度. 对 $A \in \mathscr{F}$, 定义 $\mu^+(A) := \sup\{\mu(B) : B \subset A, B \in \mathscr{F}\}$, $\mu^- := (-\mu)^+$, 那么 μ^+, μ^- 都

是 $\mathscr{F}$ 上集函数.

引理 1.1.2 定义如上, 我们有下列性质:

(1) 正集的可列并是正集;

(2) A 是正集当且仅当 $\mu^-(A)=0$; 是负集当且仅当 $\mu^+(A)=0$;

(3) μ^+, μ^- 有单调性.

证明留作习题. 下面是 Hahn 分解.

定理 1.1.3 (Hahn) 存在 $H\in\mathscr{F}$ 使得 H 是正集, 而 H^c 是负集. 集 H 称为是 μ 的 Hahn 集, 它不唯一, 可以相差一个零测集.

证明 不妨设 μ 不取到 $+\infty$. 令 $b:=\sup\{\mu(A):A\in\mathscr{F}\}$. 现在我们验证上确界可在 $\mathscr{F}$ 中达到. 首先验证对任何 $k<b$, 存在正集 $A\in\mathscr{F}$ 使得 $\mu(A)\geqslant k$. 事实上由 b 的定义, 存在 $B\in\mathscr{F}$ 使得 $\mu(B)>k$. 如果 B 是正集, 记其为 A; 如果不是, 那么对任何 $0<a_1<\mu^-(B)$, 存在 $E_1\subset B$ 使得 $\mu(E_1)<-a_1$. 如果 $B\setminus E_1$ 是正集, 记其为 A; 如果不是, 那么对任何 $0<a_2<\mu^-(B\setminus E_1)$, 存在 $E_2\subset B\setminus E_1$ 使得 $\mu(E_2)<-a_2$. 这样继续, 或者在有限步得到正集 A, 这时显然 $\mu(A)>k$, 或者得到 B 的不相交可测子集列 $\{E_n\}$ 及正数列 $\{a_n\}$ 使得 $\mu(E_n)<-a_n$,

$$a_n<\mu^-\left(B\setminus(\bigcup_{i=1}^{n-1}E_i)\right),\quad n\geqslant 1.$$

这时令 $E:=\bigcup_n E_n, A:=B\setminus E$, 则 $\mu(E)\leqslant 0$, $\mu(A)=\mu(B)-\mu(E)\geqslant\mu(B)>k$. 我们来验证 $\mu^-(A)=0$. 用反证法, 假设 $\mu^-(A)>0$. 因 $\mu^-[B\setminus(\bigcup_{i=1}^{n-1}E_i)]\geqslant\mu^-(A)$, 则由 a_n 的取法, 存在 $\lambda>0$, 使得 $a_n\geqslant\lambda$. 因此 $\mu(E)\leqslant-\sum_n a_n=-\infty$, 推出 $k<\mu(B)=\mu(A)+\mu(E)=-\infty$. 矛盾. 这样证明了 $\mu^-(A)=0$.

因此, 可取得可测正集列 $\{A_n\}$, 使得 $\lim_n\mu(A_n)=b$, 令 $H:=\bigcup_n A_n$, 则 H 是一个 Hahn 集. 事实上, H 显然是正集, 且 $\mu(H)\geqslant b$, 故 $\mu(H)=b$. 因此 $b<+\infty$, 另外对 H^c 的任何可测子集 A, $b\geqslant\mu(H\cup A)=b+\mu(A)$, 推出 $\mu(A)\leqslant 0$. 这证明了 H^c 是负集. □

由此, 我们得到 Jordan 分解.

定理 1.1.4 (Jordan) 设 H 是 Hahn 集, 则

(1) $\mu^+(A) = \mu(A \cap H)$, $\mu^-(A) = -\mu(A \cap H^c)$, $A \in \mathscr{F}$, 因此 μ^+, μ^- 是测度, 且其中至少有一个是有限的;

(2) $\mu = \mu^+ - \mu^-$ 且如果存在测度 μ_1, μ_2 使得 $\mu = \mu_1 - \mu_2$, 则 $\mu^+ \leqslant \mu_1$, $\mu^- \leqslant \mu_2$.

记 $|\mu| := \mu^+ + \mu^-$, 它称为是 μ 的全变差测度. $|\mu|(\Omega)$ 是 μ 的全变差.

习　题

1. 若 $\mathscr{F}$ 是有限多个元素的 σ 代数, 那么 $|\mathscr{F}|$ 是 2 的整数幂.
2. 设 $\{\mathscr{F}_n\}$ 是 Ω 上子集类的递增序列, 证明: 如果所有 $\mathscr{F}_n$ 是代数, 则 $\bigcup\limits_n \mathscr{F}_n$ 也是代数. 举例说明即使所有 $\mathscr{F}_n$ 是 σ- 代数, 它们的并也不一定是 σ- 代数. 如果 $\{\mathscr{F}_n\}$ 是严格递增的 σ- 代数列, 那么 $\bigcup_n \mathscr{F}_n$ 一定不是 σ- 代数.
3. 证明: $\mathscr{B}(\mathbf{R})$ 是可列生成的. 因此 $\mathbf{R}$ 上存在 Lebesgue 可测集不是 Borel 可测的.
4. 取 Ω 的子集列 $\{A_n\} \subset \mathscr{F}$. 证明: 上极限集合 $\limsup_n A_n := \bigcap_n \bigcup_{k \geqslant n} A_k$ 是属于无穷多 A_n 的元素全体, 下极限集合 $\liminf_n A_n := \bigcup_n \bigcap_{k \geqslant n} A_k$ 是属于除有限多外全体 A_n 的元素全体.
5. 设 $\mathscr{F}_0$ 是 Ω 上一个代数, μ 是 $\mathscr{F}_0$ 上有限可加集函数且 $\mu(\Omega) < +\infty$. 证明: 如果 μ 在 $\varnothing$ 处上连续, 则 μ 是预测度.
6. 证明: (1) $\mathscr{F}^\mu = \{A \cup N : A \in \mathscr{F}, N \in \mathscr{N}\}$;

   (2) $(\Omega, \mathscr{F}^\mu, \mu)$ 是一个完备测度空间. (参见定义 1.1.2 下关于完备测度空间的论述.)
7. 设 $(\Omega, \mathscr{F}, \mu)$ 是 σ- 有限测度空间, μ^* 是由 μ 诱导的外测度, 对 $A \subset \Omega$, 定义

$$\mu^{**}(A) := \inf\{\mu(B) : A \subset B, B \in \mathscr{F}\},$$

$$\mu_*(A) := \sup\{\mu(B) : A \supset B, B \in \mathscr{F}\},$$

   其中 μ_* 称为是 A 的(由 μ 诱导的)内测度, 证明:

   (1) $\mu^* = \mu^{**}$;

   (2) 记 $\mathscr{M}$ 为 μ^*- 可测集全体, 则 $(\Omega, \mathscr{M}, \mu^*)$ 是 $(\Omega, \mathscr{F}, \mu)$ 的完备化.

   (3) 如果 $\mu^*(\Omega) < \infty$, 则 $A \in \mathscr{M}$ 当且仅当 $\mu^*(A) = \mu_*(A)$.

8. (单调类定理) 设 Ω 是一非空集, 一个子集类称为是单调类, 如果它对递增(或递减)集列的极限封闭. 证明: 若 $\mathscr{A}$ 是代数, $\mathscr{F}$ 是单调类, $\mathscr{A}\subset\mathscr{F}$, 则 $\sigma(\mathscr{A})\subset\mathscr{F}$.
9. 设 $\Omega=\{a,b,c,d\}$, $\mathscr{A}=\big\{\varnothing,\Omega,\{a,b\},\{c,d\},\{a,d\},\{b,c\}\big\}$. 验证 $\mathscr{A}$ 是 Dynkin 系, 但不是 π- 类.
10. 证明: 一个 σ- 代数的势或者是有限或者是不可列的.
11. 证明: 子集类 $\mathscr{A}$ 是 Dynkin 系当且仅当 $\Omega\in\mathscr{A}$ 且 $\mathscr{A}$ 对包含差及递增集列极限运算封闭.
12. 设 Ω 是一个拓扑空间, $\mathscr{A}(\Omega)$ 与 $\mathscr{B}(\Omega)$ 分别是 Ω 上使得所有连续函数可测的最小 σ- 代数与 Borel 代数, 证明:

    (1) $\mathscr{A}(\Omega)\subset\mathscr{B}(\Omega)$;

    (2) 如果 Ω 是度量空间, 则 $\mathscr{A}(\Omega)=\mathscr{B}(\Omega)$.
13. 设 Ω 是一非空集, 一个子集类 A 称为是半环, 如果 $\varnothing\in A$, 且对任何 $A,B\in A$, 有 $A\cap B\in A$ 及 $A\setminus B$ 可表为 A 中元素的不交并. 证明:

    (1) 包含半环 A 的最小环是 A 中元素的有限不交并全体;

    (2) 设 μ 是半环上的可列可加集函数且 $\mu(\varnothing)$=0, 则 μ 可唯一地延拓为包含 A 的最小环上的预测度.
14. 设 Ω,Ω' 是两个非空集合, ξ 是 Ω 到 Ω' 的一个映射, 设 $\mathscr{A}'$ 是 Ω' 的一个子集类. 证明: $\sigma[f^{-1}(\mathscr{A}')]=f^{-1}[\sigma(\mathscr{A}')]$.
15. 设 λ 是 $[0,2\pi]$ 上的 Lebesgue 测度. 令 $\xi:x\mapsto\sin x,\ x\in[0,2\pi]$. 写出 λ 在 ξ 下在 $(\mathbf{R},\mathscr{B})$ 上的像测度.
16. 设 μ 是 $(\mathbf{R},\mathscr{B})$ 上 σ- 有限测度. 对任何 $x\in\mathbf{R}$, 定义 $\mu_x(A):=\mu(A+x)$, $A\in\mathscr{B}$, 称为是 μ 的 x- 平移. 证明: 如果 μ 平移不变, 即对任何 x, $\mu_x=\mu$, 则 μ 是 Lebesgue 测度的常数倍. 问如果 μ 仅对有理数平移不变, 结论是否成立?
17. 设 λ 是 $([0,1],\mathscr{B}[0,1])$ 上 Lebesgue 测度, $\mathbb{P}$ 是其上概率测度, 如果对任何 $A\in\mathscr{B}[0,1]$, $\lambda(A)=\dfrac{1}{2}$ 蕴含有 $\mathbb{P}(A)=\dfrac{1}{2}$, 证明: $\mathbb{P}=\lambda$.
18. 设 μ,ν 是可测空间 $(\Omega,\mathscr{F})$ 上有限测度. 证明:

    (1) 存在 $D\in\mathscr{F}$, 使得 $1_D\cdot\mu\leqslant 1_D\cdot\nu$, $1_{D^c}\cdot\mu\geqslant 1_{D^c}\cdot\nu$;

    (2) 存在唯一的测度, 记为 $\mu\wedge\nu$, 使得: (a) $\mu\wedge\nu\leqslant\mu,\nu$; (b) 如果有测度 $\kappa\leqslant\mu,\nu$, 则 $\kappa\leqslant\mu\wedge\nu$.

19. 设 μ 是 $[0,1]$ 上 Radon 测度. 证明:

    (1) $C[0,1]$ 在 $L^2(\mu)$ 中稠;

    (2) 形如 $a_0 1_{\{0\}} + \sum_{i=1}^n a_i 1_{[t_{i-1},t_i]}$, $0 = t_0 < t_1 < \cdots, t_n$, $a_0, a_1, \cdots, a_n \in \mathbf{R}$, $n \geqslant 1$ 的函数在 $L^2(\mu)$ 中稠.

20. 设 $\{\xi_i : i \in I\}$ 是 Ω 上的一族函数, $\mathscr{F} := \sigma(\{\xi_i : i \in I\})$, 证明: 如果 $A \in \mathscr{F}$, 那么存在 I 的一个可列子集 J 使得 $A \in \sigma(\{\xi_i : i \in J\})$.

21. (Vitali-Hahn-Saks) 设 $(\Omega, \mathscr{F})$ 是可测空间, $\mathbb{P}_n$ 是其上概率测度列, 如果对任何 $A \in \mathscr{F}$, $\mathbb{P}_n(A)$ 收敛, 记极限为 $\mathbb{P}(A)$, 那么 $\mathbb{P}$ 也是一个概率测度. 另外, 如果 $B_k \in \mathscr{F}$ 是一个单调下降趋于空集的集列, 那么 $\sup\limits_n \mathbb{P}_n(B_k) \downarrow 0$.

22. 可测空间 $(\Omega, \mathscr{F})$ 上测度称为 s- 有限, 如果它可以表示为可数多有限测度的和. 证明: σ- 有限测度是 s- 有限的. 举例说明 s- 有限测度不一定是 σ- 有限测度.

§1.2 可测函数与积分

在本节, 设 $(\Omega, \mathscr{F}, \mu)$ 是给定的测度空间, $(\overline{\mathbf{R}}, \overline{\mathscr{B}})$ 是广义实值可测空间(参考实分析中的定义). $(\Omega, \mathscr{F})$ 上的可测函数是指 $(\Omega, \mathscr{F})$ 到 $(\overline{\mathbf{R}}, \overline{\mathscr{B}})$ 的可测映射. Ω 上的一个可测函数 f 称为是简单的, 如果 f 仅取有限多个值, 即存在互不相同的常数 $a_1, \cdots, a_n \in \mathbf{R}$, 使得

$$f(\omega) = \sum_{i=1}^n a_i 1_{\{f=a_i\}}(\omega), \quad \omega \in \Omega.$$

如不计次序, 此表达式是唯一的. 这时我们定义

$$\mu(f) := \sum_{i=1}^n a_i \mu(\{f = a_i\}).$$

注意我们总是约定 $0 \cdot \infty = 0$. 用 $\mathbf{S}^+$ 表示 Ω 上的非负简单函数全体. 不难验证, 映射 $\mu : \mathbf{S}^+ \to \overline{\mathbf{R}}^+$ 是单调的且线性的. 对 Ω 上的任意非负可测函数 f, 定义

$$\mu(f) := \sup\{\mu(g) : 0 \leqslant g \leqslant f,\ g \in \mathbf{S}^+\},$$

称为是 f 关于 μ 的积分.

定义 1.2.1 设 f 是 Ω 上可测函数, 记 f^+, f^- 分别是 f 的正部和负部, 当 $\mu(f^+)$, $\mu(f^-)$ 两者至少有一个是有限时, 称 f 关于 μ 的积分存在, 且记 f 关于 μ 的积分为 $\mu(f) := \mu(f^+) - \mu(f^-)$; 而当 $\mu(f^+), \mu(f^-)$ 两者都有限时, 称 f 关于 μ 是可积的.

显然, 改变可测函数 f 在一个 μ- 零测集上的值不改变积分的值, 另外如果 f 非负且在一个 μ- 正测度集上等于 $+\infty$, 则 $\mu(f) = +\infty$. 关于积分, 其他常用的记号还有 $\int_\Omega f(\omega)\mu(\mathrm{d}\omega)$, $\int_\Omega f\mathrm{d}\mu$, $\int f$, $\langle f, \mu\rangle$ 等. 另外 f 在可测集 $A \in \mathscr{F}$ 上的积分定义为 $\mu(f1_A)$, 常写为 $\int_A f\mathrm{d}\mu$. 在 Euclid 空间上可测函数关于 Lebesgue 测度的积分称为是 Lebesgue 积分, 一个可测函数如果同时 Riemann 可积并 Lebesgue 可积, 则两个积分相等. Lebesgue 积分通常认为是 Riemann 积分的推广, 但在严格的意义下, 并非如此, 例如:

例 1.2.1 考虑函数 $f(x) = \dfrac{\sin x}{x}$, $x \in \mathbf{R}$. 设 λ 是 $\mathbf{R}$ 上 Lebesgue 测度, 则 $\lambda(f^+) = \lambda(f^-) = +\infty$, 故 f 的 Lebesgue 积分不存在, 而 f 是 Riemann 可积的, 且

$$\int_{-\infty}^{+\infty} \frac{\sin x}{x}\mathrm{d}x = \pi. \qquad \blacksquare$$

从定义直接可以推出的性质是积分的单调性: 即如果 $f_1 \leqslant f_2$ 是非负可测函数, 那么 $\mu(f_1) \leqslant \mu(f_2)$. 下面所谓的单调收敛定理是积分理论中最基本的也是最重要的定理.

定理 1.2.1 设 $\{f_n\}$ 是一个递增收敛于 f 的非负可测函数序列, 则

$$\mu(f) = \uparrow \lim_n \mu(f_n),$$

这里 $\uparrow \lim$ 表示极限是一个递增极限.

证明 由单调性, $\{\mu(f_n)\}$ 是一个单调增加的数列, 且

$$\mu(f) \geqslant \lim_n \mu(f_n).$$

反之, 任取一个被 f 控制的非负简单函数 g 及 $0 < \lambda < 1$, 令 $A_n := \{f_n \geqslant \lambda g\}$. 因为在 $\{f > 0\}$ 上, 有 $f > \lambda g$, 故 $A_n \uparrow \Omega$.

$$\mu(f_n) \geqslant \mu(f_n 1_{A_n}) \geqslant \lambda\mu(g1_{A_n}).$$

因 g 是简单的, 故

$$\lim_n \mu(f_n) \geqslant \lim_n \lambda\mu(g1_{A_n}) = \lambda\mu(g),$$

最后的等号利用 μ 在 $\mathbf{S}^+$ 上的线性性及下连续性. 因 λ 是任意的, 推出

$$\lim_n \mu(f_n) \geqslant \mu(g).$$

由 $\mu(f)$ 的定义推出 $\lim_n \mu(f_n) \geqslant \mu(f)$. 完成了证明. □

一个零测集外成立的性质称为几乎处处成立. 比如, 测度空间 $(\Omega, \mathscr{F}, \mu)$ 上可测函数 f_1 与 f_2 称为几乎处处相等, 是指它们在一个 μ- 零测集外相等, 记为 $f_1 = f_2$, μ- a.e. 在上下文明确时, 简写为 $f_1 = f_2$ a.e. 或 $f_1 = f_2$. 上面定理的单调性可以用几乎处处单调代替. 另外, 任何非负可测函数 f 都可以表示为一个单调上升的非负简单可测函数序列的极限, 如

$$f = \uparrow \lim_{n\to\infty} \left(\sum_{k=1}^{n2^n} \frac{k-1}{2^n} 1_{\{\frac{k-1}{2^n} \leqslant f < \frac{k}{2^n}\}} + n1_{\{f \geqslant n\}} \right),$$

因此由单调收敛定理可以推出其积分是非负简单可测函数序列的积分的单调上升极限, 因而积分的性质通常只需对非负简单可测函数验证. 比如用单调收敛定理容易验证对任何非负可测函数 f, g 有 $\mu(f+g) = \mu(f) + \mu(g)$. 然后, 如果 f, g 可积, 那么 $f+g$ 也可积且因为 $(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+$, 故由可积性与积分定义推出

$$\mu(f+g) = \mu(f) + \mu(g).$$

读者可自行验证其单调性与其他一些简单性质. 利用这个思想也容易证明下面的关于积分的变量替换公式.

定理 1.2.2 设 f 是可测函数, ϕ 是 $\mathbf{R}$ 上 Borel 可测函数, 则 ϕ 在 $\mu \circ f^{-1}$ 下可积当且仅当 $\phi \circ f$ 在 μ 下可积, 且这时有

$$\mu(\phi \circ f) = \mu \circ f^{-1}(\phi).$$

证明 公式显然对 $\phi = 1_A$, $A \in \mathscr{B}$ 成立, 因为这就是像测度的定义. 因此公式对 $\phi \in \mathbf{S}^+$ 成立, 然后运用单调收敛定理, 对非负可测函数成立. 因而公式对使公式两

边积分存在的 ϕ 成立. □

下面我们将介绍的 Fatou 引理和 Lebesgue 控制收敛定理是分析中最重要的工具之一. 先介绍 Fatou 引理.

定理 1.2.3 (Fatou)　设 $\{f_n\}$ 是非负可测函数序列, 则

$$\mu(\liminf_n f_n) \leqslant \liminf_n \mu(f_n).$$

证明　令 $g_n := \inf_{k \geqslant n} f_k$, 则 $\{g_n\}$ 是一个单调增加的非负可测函数序列且 $g_n \leqslant f_n$, 由单调收敛定理,

$$\mu(\liminf_n f_n) = \mu(\lim_n g_n) = \lim_n \mu(g_n) = \liminf_n \mu(g_n) \leqslant \liminf_n \mu(f_n).$$

□

例 1.2.2　设 λ 是 $[0,1]$ 上 Lebesgue 测度. 定义

$$f_{2n-1} := 1_{[0,\frac{1}{2})},\ f_{2n} := 1_{(\frac{1}{2},1]},\ n \geqslant 1,$$

则 $\liminf f_n = 0$, 而 $\liminf \lambda(f_n) = \dfrac{1}{2}$.

下面是 Lebesgue 控制收敛定理.

定理 1.2.4 (Lebesgue)　设 $\{f_n\}$ 是 Ω 上的可测函数序列, 如果对任何 $\omega \in \Omega$, $\{f_n(\omega)\}$ 收敛且存在一个关于 μ 可积的可测函数 g 满足 $|f_n| \leqslant g$, 则 $\mu(\lim f_n) = \lim \mu(f_n)$.

证明　记 $f := \lim_n f_n$. 因为 $\{g - f_n\}$ 与 $\{g + f_n\}$ 都是非负可测函数序列, 利用 Fatou 引理,

$$\mu(g+f) = \mu[\liminf(g+f_n)] \leqslant \liminf \mu(g+f_n) = \mu(g) + \liminf \mu(f_n).$$

因 g 可积, 故而 $\mu(f) \leqslant \liminf \mu(f_n)$. 同理对 $\{g - f_n\}$ 用 Fatou 引理, 有 $\limsup \mu(f_n) \leqslant \mu(f)$, 由此得 $\mu(f) = \lim \mu(f_n)$. □

自然数集上有一个计数测度, 一个数列 $a = (a_n)$ 自然地看成为自然数集上的函数, 那么 a 关于计数测度的积分就是级数 $\sum\limits_{n \geqslant 1} a_n$. 因此 Lebesgue 控制收敛定理经常用于级数的求和与极限交换的问题.

推论 1.2.1　设 $\{a_{n,m}\}$ 是一个两指标的数列. 如果

(1) 对任何 n, $\lim_m a_{n,m}$ 存在;

(2) 存在数列 $\{b_n\}$ 使得 $\sum\limits_n b_n$ 收敛且对任何 m, $|a_{n,m}| \leqslant b_n$.

则

$$\lim_m \sum_n a_{n,m} = \sum_n \lim_m a_{n,m}.$$

在有限测度空间特别是在概率空间上, 有界可测函数总是可积的, 因而我们有下面的有界收敛定理.

推论 1.2.2 设 $(\Omega, \mathscr{F}, \mu)$ 是一个有限测度空间, $\{f_n\}$ 是 Ω 上的可测函数序列, 如果对任何 $\omega \in \Omega$, $\{f_n(\omega)\}$ 收敛且存在一个常数 M 使得 $|f_n| \leqslant M$, 则 $\mu(\lim\limits_n f_n) = \lim\limits_n \mu(f_n)$.

例 1.2.3 此例说明 Lebesgue 控制收敛定理中的条件如不满足, 结论未必成立. 设 λ 是 $[0,1]$ 上 Lebesgue 测度, 定义 $f_n := n \cdot 1_{(0,\frac{1}{n}]}$, 则 f_n 点点收敛于函数 0, 但 $\lambda(f_n)$ 恒等于 1.

实际上, 在后面引入几乎处处收敛的概念后, Lebesgue 控制收敛定理和有界收敛定理中的处处收敛的条件可由几乎处处收敛代替. 下面我们讨论测度的绝对连续和 Radon-Nikodym 导数.

设 μ 是可测空间 $(\Omega, \mathscr{F})$ 上测度, f 是非负可测函数, 定义 $\nu(A) := \mu(f \cdot 1_A)$, $A \in \mathscr{F}$. 则 ν 也是 $(\Omega, \mathscr{F})$ 上测度, 记 ν 为 $f \cdot \mu$. 对 $A \in \mathscr{F}$, 测度 $1_A \cdot \mu$ 实际上是 μ 在 A 上的限制. 任给两个测度 μ, ν, 如果存在一个可测函数 f, 使得 $\nu = f \cdot \mu$, 那么说 ν 关于 μ 是 (Radon-Nikodym) 可导的, 并称 f 是 ν 关于 μ 的 Radon-Nikodym 导数, 常写为 $\dfrac{\mathrm{d}\nu}{\mathrm{d}\mu}$. 另外称 ν 关于 μ 绝对连续, 如果 $A \in \mathscr{F}$ 且 $\mu(A) = 0$ 蕴含着 $\nu(A) = 0$. 记为 $\nu \ll \mu$. 显然如果 ν 关于 μ 是 Radon-Nikodym 可导的, 则 $\nu \ll \mu$, 但一般地反之不对, 如奇异测度与计数测度. 但若 μ 是 σ- 有限时, 逆命题成立.

定理 1.2.5 (Radon-Nikodym) 设 μ, ν 是可测空间 $(\Omega, \mathscr{F})$ 上两个 σ- 有限的测度. 如果 $\nu \ll \mu$, 则 ν 关于 μ 可导, 其 Radon-Nikodym 导数 μ- 几乎处处有限且在 μ- 几乎处处相等的意义下是唯一的.

证明 在这里, 让我们仅证明 Radon-Nikodym 导数存在. 不妨设 ν, μ 是有限测度, 令

$$\mathscr{H} := \{h \geqslant 0: \text{ 关于 } \mathscr{F} \text{ 可测且 } h \cdot \mu \leqslant \nu\},$$

$b := \sup\{\mu(h) : h \in \mathscr{H}\}$. 我们先证明 $\mathscr{H}$ 对运算 $\vee$ 封闭. 事实上, 取 $h_1, h_2 \in \mathscr{H}$, 记 $h := h_1 \vee h_2$, 则

$$\begin{aligned} h \cdot \mu &= h 1_{\{h_1 \leqslant h_2\}} \cdot \mu + h 1_{\{h_1 > h_2\}} \cdot \mu \\ &= h_2 1_{\{h_1 \leqslant h_2\}} \cdot \mu + h_1 1_{\{h_1 > h_2\}} \cdot \mu \\ &\leqslant 1_{\{h_1 \leqslant h_2\}} \cdot \nu + 1_{\{h_1 > h_2\}} \cdot \nu = \nu, \end{aligned}$$

因此 $h \in \mathscr{H}$. 因此可以取递增函数列 $\{h_n\} \subset \mathscr{H}$, 使得 $b = \lim \mu(h_n)$. 令 $g := \lim h_n$, 则由单调收敛定理有 $\mu(g) = b$ 且 $g \in \mathscr{H}$.

现在我们将证明 $\nu = g \cdot \mu$. 定义 $\lambda := \nu - g \cdot \mu$. 因为 $g \in \mathscr{H}$, 故 λ 是一个有限测度. 对任何 $n \geqslant 1$, 由 Hahn 分解, 存在 $D_n \in \mathscr{F}$, 使得 $1_{D_n} \cdot \lambda \leqslant 1_{D_n} \cdot \dfrac{1}{n}\mu$ 及 $1_{D_n^c} \cdot \lambda \geqslant 1_{D_n^c} \cdot \dfrac{1}{n}\mu$. 推出 $g + \dfrac{1}{n} 1_{D_n^c} \in \mathscr{H}$. 而 g 的最大性蕴含着 $\mu(D_n^c) = 0$. 令 $D := \bigcap_n D_n$, 则 $\mu(D^c) = 0$, 由绝对连续性得 $\lambda(D^c) = 0$. 另一方面, 对任何 n, $\lambda(D) \leqslant \lambda(D_n) \leqslant \dfrac{1}{n}\mu(D_n) \leqslant \dfrac{1}{n}\mu(\Omega)$, 因此有 $\lambda(D) = 0$. □

在本节最后我们介绍乘积测度空间与 Fubini 定理.

设 $(\Omega_1, \mathscr{F}_1, \mu_1)$ 和 $(\Omega_2, \mathscr{F}_2, \mu_2)$ 是两个 σ- 有限测度空间, 记

$$\mathscr{F}_1 \otimes \mathscr{F}_2 := \{A_1 \times A_2 : \ A_1 \in \mathscr{F}_1, \ A_2 \in \mathscr{F}_2\},$$

则 $\mathscr{F}_1 \otimes \mathscr{F}_2$ 是乘积空间 $\Omega_1 \times \Omega_2$ 上的一个 π- 类. 令 $\mathscr{F}_1 \times \mathscr{F}_2 := \sigma(\mathscr{F}_1 \otimes \mathscr{F}_2)$, 称为是乘积 σ- 代数. 对任意 $Q \subset \Omega_1 \times \Omega_2$ 和 $\omega_1 \in \Omega_1$, 用 Q^{ω_1} 表示 Q 在 Ω_2 上的 ω_1-截面, 即

$$Q^{\omega_1} := \{\omega_2 \in \Omega_2 : \ (\omega_1, \omega_2) \in Q\}.$$

同理对 $\omega_2 \in \Omega_2$, 用 Q_{ω_2} 表示 Q 在 Ω_1 上的 ω_2- 截面. 容易验证, 如果 $\omega_1 \in A_1$, 则 $(A_1 \times A_2)^{\omega_1} = A_2$, 否则 $(A_1 \times A_2)^{\omega_1} = \varnothing$.

引理 1.2.1　设 $Q \in \mathscr{F}_1 \times \mathscr{F}_2$, 则

(1) 对 $\omega_1 \in \Omega_1$, $Q^{\omega_1} \in \mathscr{F}_2$; 对 $\omega_2 \in \Omega_2$, $Q_{\omega_2} \in \mathscr{F}_1$.

(2) $\omega_1 \mapsto \mu_2(Q^{\omega_1})$ 是 Ω_1 上可测函数, 同样 $\omega_2 \mapsto \mu_1(Q_{\omega_2})$ 是 Ω_2 上的可测函数.

证明　(1) 令

$$\mathscr{A} := \{A \subset \Omega_1 \times \Omega_2 : \ A^{\omega_1} \in \mathscr{F}_2\}.$$

容易验证 $\mathscr{A}$ 是 $\Omega_1 \times \Omega_2$ 上的 σ- 代数, 且 $\mathscr{A} \supset \mathscr{F}_1 \otimes \mathscr{F}_2$, 因此 $\mathscr{A} \supset \mathscr{F}_1 \times \mathscr{F}_2$.

(2) 同样令

$$\mathscr{A} := \{A \subset \Omega_1 \times \Omega_2 : \ \omega_1 \mapsto \mu_2(A^{\omega_1}) \text{ 是 } \Omega_1 \text{ 上可测函数}\}.$$

容易验证对任何 $A_1 \in \mathscr{F}_1, A_2 \in \mathscr{F}_2$, 有 $\mu_2((A_1 \times A_2)^{\omega_1}) = \mu_2(A_2) \cdot 1_{A_1}(\omega_1)$, 故 $\mathscr{A} \supset \mathscr{F}_1 \otimes \mathscr{F}_2$. 另外不难验证 $\mathscr{A}$ 是一个 λ- 类, 但是 $\mathscr{F}_1 \otimes \mathscr{F}_2$ 是 π- 类, 由 Dynkin 的 λ- π 方法知 $\mathscr{A} \supset \sigma(\mathscr{F}_1 \otimes \mathscr{F}_2) = \mathscr{F}_1 \times \mathscr{F}_2$. □

对 $\omega_1 \in \Omega_1$, 定义 $T^{\omega_1}(\omega_2) = (\omega_1, \omega_2), \omega_2 \in \Omega_2$, 类似地定义 $T_{\omega_2} : \Omega_1 \to \Omega_2 \times \Omega_2$. 显然, $(T^{\omega_1})^{-1}(Q) = Q^{\omega_1}$, $Q \in \Omega_1 \times \Omega_2$. 上述引理中 (1) 等价于说 $T^{\omega_1}, T_{\omega_2}$ 是可测映射. 故对任何 $\Omega_1 \times \Omega_2$ 上可测函数 f, 固定 ω_1, $f(\omega_1, \cdot) = f \circ T^{\omega_1}$ 是 Ω_2 上可测函数.

再由引理中 (2), 我们将定义

$$\mu_1 \times \mu_2(Q) := \int_{\Omega_1} \mu_2(Q^{\omega_1})\mu_1(\mathrm{d}\omega_1), \ Q \in \mathscr{F}_1 \times \mathscr{F}_2.$$

容易验证 $\mu_1 \times \mu_2$ 是 $(\Omega_1 \times \Omega_2, \mathscr{F}_1 \times \mathscr{F}_2)$ 上测度. 而且因 μ_1, μ_2 是 σ- 有限的, 故 $\mu_1 \times \mu_2$ 也是 σ- 有限的. 另一方面, 对 $A_1 \in \mathscr{F}_1, A_2 \in \mathscr{F}_2$,

$$\mu_1 \times \mu_2(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2) = \int_{\Omega_2} \mu_1((A_1 \times A_2)_{\omega_2})\mu_2(\mathrm{d}\omega_2).$$

因 $\mathscr{F}_1 \otimes \mathscr{F}_2$ 是 π- 类, 扩张唯一性定理蕴含着对任何 $Q \in \mathscr{F}_1 \times \mathscr{F}_2$, 有

$$\mu_1 \times \mu_2(Q) = \int_{\Omega_1} \mu_2(Q^{\omega_1})\mu_1(\mathrm{d}\omega_1) = \int_{\Omega_2} \mu_1(Q_{\omega_2})\mu_2(\mathrm{d}\omega_2). \tag{1.2.1}$$

任取 $\Omega_1 \times \Omega_2$ 上非负可测函数 f, 当 f 是简单的时, 由积分的线性性质,

$$\omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \omega_2)\mu_2(\mathrm{d}\omega_2)$$

是可测函数, 而一般的非负可测函数是非负简单函数的递增极限, 由单调收敛定理, 上面的函数当 f 是一般非负可测函数时是 Ω_1 上可测函数. 同样如果 f 是简单的,

那么由公式 (1.2.1) 及积分的线性性质, 有

$$
\begin{aligned}
\int_{\Omega_1\times\Omega_2} f\mathrm{d}\mu_1\times\mu_2 &= \int_{\Omega_2}\mu_2(\mathrm{d}\omega_2)\int_{\Omega_1} f(\omega_1,\omega_2)\mu_1(\mathrm{d}\omega_1)\\
&= \int_{\Omega_1}\mu_1(\mathrm{d}\omega_1)\int_{\Omega_2} f(\omega_1,\omega_2)\mu_2(\mathrm{d}\omega_2).
\end{aligned}
\tag{1.2.2}
$$

类似地, 单调收敛定理推出上面的公式对一般非负可测函数 f 成立.

再考虑可积的 f. 这时对它的正部分 f^+ 应用公式 (1.2.2), 有

$$
\int_{\Omega_1}\mu_1(\mathrm{d}\omega_1)\int_{\Omega_2} f^+(\omega_1,\omega_2)\mu_2(\mathrm{d}\omega_2) = \int_{\Omega_1\times\Omega_2} f^+\mathrm{d}\mu_1\times\mu_2 < +\infty.
$$

因此, 不难看出 $\omega_1\mapsto\int_{\Omega_2} f^+(\omega_1,\omega_2)\mu_2(\mathrm{d}\omega_2)$ 是关于 μ_1 几乎处处有限并是可积的, 从而

$$
\begin{aligned}
\omega_1\mapsto &\int_{\Omega_2} f(\omega_1,\omega_2)\mu_2(\mathrm{d}\omega_2)\\
&= \int_{\Omega_2} f^+(\omega_1,\omega_2)\mu_2(\mathrm{d}\omega_2) - \int_{\Omega_2} f^-(\omega_1,\omega_2)\mu_2(\mathrm{d}\omega_2)
\end{aligned}
$$

也是关于 μ_1 几乎处处有定义(取实值)并是可积的. 再应用上面非负函数的公式 (1.2.2), 简略地写, 有

$$
\begin{aligned}
\int_{\Omega_1\times\Omega_2} f\mathrm{d}\mu_1\times\mu_2 &= \int f^+\mathrm{d}\mu_1\times\mu_2 - \int f^-\mathrm{d}\mu_1\times\mu_2\\
&= \int \mathrm{d}\mu_1\int f^+\mathrm{d}\mu_2 - \int\mathrm{d}\mu_1\int f^-\mathrm{d}\mu_2\\
&= \int \mathrm{d}\mu_1\left(\int f^+\mathrm{d}\mu_2 - \int f^-\mathrm{d}\mu_2\right)\\
&= \int \mathrm{d}\mu_1\int f\mathrm{d}\mu_2.
\end{aligned}
$$

同理可证明另一个等式.

这样我们实际上已经证明了 Fubini 的积分序交换公式.

定理 1.2.6 (Fubini) 设 $(\Omega_1,\mathscr{F}_1,\mu_1)$ 和 $(\Omega_2,\mathscr{F}_2,\mu_2)$ 是两个 σ- 有限测度空间, f 是 $(\Omega_1\times\Omega_2,\mathscr{F}_1\times\mathscr{F}_2)$ 上的可测函数. 如果 f 是非负的或者可积的, 则二重积分等

于累次积分

$$\int_{\Omega_1\times\Omega_2} f\mathrm{d}\mu_1\times\mu_2 = \int_{\Omega_1}\mu_1(\mathrm{d}\omega_1)\int_{\Omega_2} f(\omega_1,\omega_2)\mu_2(\mathrm{d}\omega_2)$$
$$= \int_{\Omega_2}\mu_2(\mathrm{d}\omega_2)\int_{\Omega_1} f(\omega_1,\omega_2)\mu_1(\mathrm{d}\omega_1).$$

例 1.2.4　让我们用 Fubini 定理计算著名的积分

$$I := \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2}\mathrm{d}x.$$

事实上

$$I^2 = \int_{\mathbf{R}} \mathrm{e}^{-x^2}\mathrm{d}x\int_{\mathbf{R}} \mathrm{e}^{-y^2}\mathrm{d}y$$
$$= \int \mathrm{d}x\int \mathrm{e}^{-(x^2+y^2)}\mathrm{d}y = \int_{\mathbf{R}^2} \mathrm{e}^{-(x^2+y^2)}\mathrm{d}x\mathrm{d}y,$$

其中 $\mathrm{d}x\mathrm{d}y$ 是 $\mathbf{R}^2$ 上 Lebesgue 测度. 令 $x = r\cos\theta, y = r\sin\theta$, 则

$$I^2 = 2\pi\int_0^{\infty} r\cdot\mathrm{e}^{-r^2}\mathrm{d}r = \pi.$$

因此 $I = \sqrt{\pi}$. ▮

例 1.2.5　Fubini 定理条件中的 σ- 有限性是必需的. 设 $I = [0,1]$, μ_1,μ_2 分别是 I 上的 Lebesgue 测度与计数测度. 令

$$f(x,y) := 1_{\{x=y\}},\ x,y\in I.$$

那么容易计算

$$\int_I \mathrm{d}\mu_1\int_I f\,\mathrm{d}\mu_2 = 1,\text{ 而 }\int_I \mathrm{d}\mu_2\int_I f\,\mathrm{d}\mu_1 = 0.$$

因此 Fubini 定理不成立, 原因是计数测度不是 σ- 有限的. ▮

习　　题

1. 设 $\mathscr{H}$ 是 Ω 上的一个函数族. 令 $\sigma(\mathscr{H}) := \sigma(\{f^{-1}(B) : B\in\overline{\mathscr{B}}, f\in\mathscr{H}\})$. 证明: $\sigma(\mathscr{H})$ 是 Ω 上使得 $\mathscr{H}$ 中函数成为可测的最小 σ- 代数, 即满足:

   (1) $\mathscr{H}$ 中的函数是 $\sigma(\mathscr{H})$- 可测的;

(2) 如果 $\mathscr{F}$ 是 Ω 上 σ- 代数, 且 $\mathscr{H}$ 中的函数是 $\mathscr{F}$- 可测的, 则 $\mathscr{F} \supset \sigma(\mathscr{H})$.

2. (函数的单调类定理) 设 Ω 是一非空集, $\mathscr{A}$ 是 Ω 上一 π- 类, $\mathscr{H}$ 是 Ω 上一些实值函数构成的线性空间. 证明: 如果满足:

(1) $1 \in \mathscr{H}$;

(2) 若 $\{f_n\} \subset \mathscr{H}$, $0 \leqslant f_n \uparrow f$ 且 f 有限, 则 $f \in \mathscr{H}$;

(3) 对任何 $A \in \mathscr{A}$ 有 $1_A \in \mathscr{H}$,

则 $\mathscr{H}$ 包含 Ω 上的 $\sigma(\mathscr{A})$- 可测的实值函数.

3. 设 ξ, η 是 Ω 上两可测函数, 则 ξ 关于 $\sigma(\eta)$ 可测当且仅当存在 $\mathbf{R}$ 上 Borel 可测函数 f 使得 $\xi = f \circ \eta$.

4. 设 μ, ν 是 $(\Omega, \mathscr{F})$ 上两个测度, 称它们互相奇异 (记为 $\mu \perp \nu$), 如果存在 $N \in \mathscr{F}$ 使得 $\mu(N) = 0$, $\nu(N^c) = 0$. 证明: (Lebesgue 分解) 当 μ, ν 是有限测度时, 存在非负可测函数 g 使得 $\nu - g \cdot \mu$ 是测度且与 μ 互相奇异.

5. 设 μ, ν 是 $(\Omega, \mathscr{F})$ 上两个有限测度, 令 $\lambda := \mu + \nu$. 证明:

(1) 存在 $0 \leqslant g \leqslant 1$ a.e. λ 使得

$$\int_\Omega f \mathrm{d}\nu = \int_\Omega fg \mathrm{d}\lambda, \ f \in L^2(\lambda);$$

(2) 令 $A := \{0 \leqslant g < 1\}$, $B = \{g = 1\}$. 定义 $\nu_s := 1_B \cdot \nu$, $\nu_a := 1_A \cdot \nu$. 则 $\nu_s \perp \mu$, $\nu_a \ll \mu$ 且 $\nu = \nu_s + \nu_a$;

(3) 利用上面的结论直接证明: 存在 $h \in L^1(\mu)$ 使得 $\nu_a = h \cdot \mu$.

6. 一个概率测度空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 称为是非原子的, 如果对任何 $\mathbb{P}(A) > 0$, 存在 $B \subset A$ 使得 $0 < \mathbb{P}(B) < \mathbb{P}(A)$. (自然 $A, B \in \mathscr{F}$.) 证明:

(1) Lebesgue 测度是非原子的;

(2) 如果以上概率空间是非原子的, 那么对任何 $\mathbb{P}(A) > 0$, $\epsilon > 0$, 存在 $B \subset A$, 使得 $0 < \mathbb{P}(B) < \epsilon$;

(3) 如果以上概率空间是非原子的, 那么对任何 $0 < x < \mathbb{P}(A)$, 存在 $B \subset A$, 使得 $\mathbb{P}(B) = x$.

7. Ω 上的一些实值函数组成的线性空间 $\mathscr{H}$ 称为是一个向量格, 如果 $f \in \mathscr{H}$ 蕴含 $|f|, 1 \wedge f \in \mathscr{H}$. 如果 I 是向量格 $\mathscr{H}$ 上的正线性泛函且满足 $\{f_n\} \subset \mathscr{H}, f_n \downarrow 0$

蕴含着 $I(f_n)\downarrow 0$, 称 I 是 $\mathscr{H}$ 上 的 Daniell 积分. 证明: 如果 $(\Omega,\mathscr{F},\mu)$ 是测度空间, 则简单函数全体 $\mathscr{H}_1$ 与 $\mathscr{H}_2=L^1(\Omega,\mathscr{F},\mu)$ 是向量格, 且 $I(\cdot):=\mu(\cdot)$ 分别是 $\mathscr{H}_1$, $\mathscr{H}_2$ 上的 Daniell 积分.

8. 设 $\mathscr{H}$ 是 Ω 上向量格, I 是 $\mathscr{H}$ 上 Daniell 积分. 令 $\mathscr{F}_0$ 是 Ω 的满足条件: 存在非负 $\{f_n\}\subset\mathscr{H}$ 使得 $f_n\uparrow 1_A$, 的子集 A 全体. 定义 $\mu^*(A):=\lim_n I(f_n)$, $A\in\mathscr{F}_0$, 其中 $\{f_n\}$ 如上; $\mu^*(B):=\inf\{\mu^*(A):A\in\mathscr{F}_0, A\supset B\}$, $B\subset\Omega$. 证明:

   (Daniell-Stone) (1) $\sigma(\mathscr{F}_0)=\sigma(\mathscr{H})$;

   (2) μ^* 是 Ω 上外测度;

   (3) 用 $\mathscr{M}$ 表示 μ^*- 可测子集全体, 则 $\mathscr{M}\supset\mathscr{F}_0$. 因此, 存在 $(\Omega,\sigma(\mathscr{F}_0))$ 上的测度 μ 使得对任何 $f\in\mathscr{H}$, $\mu(f)=I(f)$.

9. 设 μ,ν 是 $(\Omega,\mathscr{F})$ 上两个测度, ν 是有限的. 证明: $\nu\ll\mu$ 当且仅当对任何 $\epsilon>0$, 存在 $\delta>0$, 使得 $\mu(A)<\delta$ 蕴含着 $\nu(A)<\epsilon$.

10. 证明: 如果 μ 是 σ- 有限的, 则存在非负可积函数 f 使得 $f\cdot\mu$ 是一个概率测度.

11. 设可测 f 关于 σ- 有限测度 μ 的积分存在, 令 $\nu:=f\cdot\mu$. 证明: ν 是一个符号测度, 且 $|\nu|=|f|\cdot\mu$.

12. 设 $\Omega=\{1,2,\cdots,n\}$, $\mathscr{F}$ 是其全体子集, $\nu(\{k\})=a_k$, $\mu(\{k\})=b_k$, $k\geqslant 1$.

    (1) 给出 $\nu\ll\mu$ 的充要条件;

    (2) 如果 $\nu\ll\mu$, 计算 $\dfrac{\mathrm{d}\nu}{\mathrm{d}\mu}$.

13. 设 $(X,\mathscr{X},\mu)=(Y,\mathscr{Y},\nu)$ 是 Lebesgue 测度空间 $(\mathbf{R},\mathscr{B},\lambda)$ 的完备化. 请说明 $(X\times Y,\mathscr{X}\times\mathscr{Y},\mu\times\nu)$ 不是完备的.

14. 验证 $\dfrac{xy}{(x^2+y^2)^2}$ 在 $\{(x,y):|x|\leqslant 1,|y|\leqslant 1\}$ 上(考虑 Lebesgue 测度的乘积)二重积分不存在, 而它的两个累次积分存在并相等.

15. 计算 Riemann 积分

$$\int_0^\infty \frac{\sin x}{x}\mathrm{d}x.$$

§1.3　随机变量与分布

定义 1.3.1　一个三元组 $(\Omega,\mathscr{F},\mathbb{P})$ 称为是一个概率空间, 如果 Ω 是一个非空集合, $\mathscr{F}$ 是 Ω 上的 σ- 代数且 $\mathbb{P}$ 是 $(\Omega,\mathscr{F})$ 上的一个概率测度. 这时候, 也称 Ω 是样本空间, $\mathscr{F}$ 为事件域, $\mathscr{F}$ 中的元素为事件, 而 $\mathbb{P}$ 是概率.

事件 Ω 称为是必然事件, 空集 $\varnothing$ 为不可能事件, 如果一个 Ω 上的性质在一个概率为零的事件之外成立, 称此性质在 Ω 上几乎处处成立, 或说其以概率 1 成立.

例 1.3.1 (古典等概率模型)　设 Ω 是一个有限非空集合, $\mathscr{F}$ 是 Ω 的子集全体, 它是 Ω 上的一个平凡 σ- 代数, 对 $A\subset\Omega$, 定义

$$\mathbb{P}(A):=\frac{\#(A)}{\#(\Omega)},$$

其中 $\#$ 表示 Ω 上的计数测度. 显然 $(\Omega,\mathscr{F},\mathbb{P})$ 是一个概率空间. ▮

例 1.3.2 (几何等概率模型)　与古典等可能性类似的是几何等可能性. 比如我们说在一个线段上随机地任取一个点, 或者在一个圆内随机地取一个点等等. 这自然隐含着指每个点取到的可能性是一样的, 但这时可能的结果是所有点全体是一个无限元素的集合, 故而等可能性不能如古典情形那样用数元素的个数来描述.

我们可以用 Lebesgue 测度的概念严格地给出等可能性的含义. 首先我们知道在一个给定线段上随机地取一个点落在区间 I 的概率只与区间的长度有关, 而与其位置无关. 也就是说概率应该是区间的长度 $|I|$ 与整个线段的长度 $|\Omega|$ 的比. 类似地, 给定一个有界区域 $\Omega\subset\mathbf{R}^d$, 在区域上随机地任取一个点, 那么这个点落在(Lebesgue 可测)子集 D 中的概率等于其 Lebesgue 测度 $|D|$ 与整个区域的测度 $|\Omega|$ 的比. 这样的概率我们称为是区域 Ω 上的均匀分布. ▮

首先让我们考虑 Euclid 空间, 空间 $(\mathbf{R}^n,\mathscr{B}^n)$ 上的一个概率测度称为是一个 n-维分布. 一个 1 维分布简称为分布. 怎样具体构造 $\mathbf{R}^n$ 上的分布呢? 让我们先引入分布函数的概念.

设 $a=(a_1,\cdots,a_n),b=(b_1,\cdots,b_n)\in\mathbf{R}^n$, 说 $a\leqslant b$ 是指 $a_1\leqslant b_1,\cdots,a_n\leqslant b_n$. 设 F 是 $\mathbf{R}^n$ 上的一个函数, 对 $a,b\in\mathbf{R}^n$, 令

$$\Delta_{a_k,b_k}F(x_1,\cdots,x_k,\cdots,x_n)$$
$$:= F(x_1,\cdots,b_k,\cdots,x_n) - F(x_1,\cdots,a_k,\cdots,x_n);$$
$$\Delta_a^b F := \Delta_{a_n,b_n}(\Delta_{a_{n-1},b_{n-1}}\cdots(\Delta_{a_1,b_1}F)\cdots).$$

例如, $n=1$ 时, $\Delta_{a_1}^{b_1}F = F(b_1)-F(a_1)$; $n=2$ 时,

$$\Delta_{(a_1,a_2)}^{(b_1,b_2)}F = F(b_1,b_2)-F(b_1,a_2)-F(a_1,b_2)+F(a_1,a_2).$$

定义 1.3.2 $\mathbf{R}^n$ 上的函数 F 称为是分布函数, 如果它满足下列条件:

(1) 对任意 $a,b\in\mathbf{R}^n$ 且 $a\leqslant b$, 则 $\Delta_a^b F\geqslant 0$;

(2) F 对于每一个分量是右连续的;

(3) 对每个 $0\leqslant k\leqslant n$, $\lim_{x_k\to-\infty}F(x_1,\cdots,x_n)=0$, 且

$$\lim_{x_1\to+\infty,\cdots,x_n\to+\infty}F(x_1,\cdots,x_n)=1.$$

$\mathbf{R}$ 上的函数 F 是分布函数当且仅当 F 是递增右连续且

$$\lim_{x\to-\infty}F(x)=0,\quad \lim_{x\to+\infty}F(x)=1.$$

如果 μ 是 $\mathbf{R}^n$ 上的一个分布, 定义函数

$$F(x_1,\cdots,x_n) := \mu((-\infty,x_1]\times\cdots\times(-\infty,x_n]).$$

不难验证 $\mu((a,b]) = \Delta_a^b F$, 其中 $a,b\in\mathbf{R}^n$, $a\leqslant b$ 且 $(a,b] := (a_1,b_1]\times\cdots\times(a_n,b_n]$, 故而容易验证 F 是 $\mathbf{R}^n$ 上的分布函数. 下面的定理说明逆命题也成立, 因此 $\mathbf{R}^n$ 上的分布全体与分布函数全体是一一对应的.

定理 1.3.1 设 F 是 $\mathbf{R}^n$ 上的分布函数, 则在 $\mathbf{R}^n$ 上存在唯一的分布 μ, 使对任意 $x\in\mathbf{R}^n$, 有

$$F(x_1,\cdots,x_n) = \mu((-\infty,x_1]\times\cdots\times(-\infty,x_n]).$$

证明 不妨设 $n=1$, 对 $-\infty<a\leqslant b<+\infty$, 定义 $\mu((a,b]) := F(b)-F(a)$, 那么 μ 可以自然地延拓到形如 $(a,b]$ 区间的有限不交并组成的集类 $\mathscr{F}_0$ 上. 显然 $\mathscr{F}_0$ 是一个环. 因此由 Carathéodory 扩张定理, 我们只需验证 μ 是 $\mathscr{F}_0$ 上的预测度.

μ 在 $\mathscr{F}_0$ 上的有限可加性是显然的, 因此只需验证其次可列可加性, 设有区间

列 $\{(a_i,b_i]:i\geqslant 1\}$ 使得 $(a,b]\subset\bigcup_{i=1}^{\infty}(a_i,b_i]$. 对任何 $\epsilon>0$, 由 F 的右连续性, 对任何 i, 存在 $\delta_i>0$ 使 $F(b+\delta_i)-F(b)\leqslant\frac{\epsilon}{2^i}$ 及 $\delta>0$ 使得 $F(a+\delta)-F(a)\leqslant\epsilon$. 而 $[a+\delta,b]\subset\bigcup_{i=1}^{\infty}(a_i,b_i+\delta_i)$, 由有限覆盖定理, 存在有限整数集 I 使得

$$(a+\delta,b]\subset[a+\delta,b]\subset\bigcup_{i\in I}(a_i,b_i+\delta_i)\subset\bigcup_{i\in I}(a_i,b_i+\delta_i].$$

由有限可加性, $F(b)-F(a+\delta)\leqslant\sum_{i\in I}F(b_i+\delta_i)-F(a_i)$, 因此

$$\mu((a,b])-\epsilon\leqslant\sum_i\mu((a_i,b_i])+\epsilon,$$

推出 μ 的次可列可加性.

最后, 我们还需要验证得到的 Carathéodory 扩张是个概率测度, 事实上, 由下连续性 $\mu(\mathbf{R})=\lim_n\mu((-n,n])=\lim_n[F(n)-F(-n)]=1$. 完成证明. □

需要时, 我们用 F_μ 表示分布 μ 对应的分布函数, 用 μ_F 表示分布函数 F 对应的分布.

一般地, 概率空间通常是在抽象的集合上定义, 不便运算. 因此在许多情况下, 我们引入随机变量, 将概率投射到 Euclid 空间上讨论. 给定概率空间 $(\Omega,\mathscr{F},\mathbb{P})$. 一个 n- 维随机变量是指 $(\Omega,\mathscr{F})$ 到 $(\mathbf{R}^n,\mathscr{B}^n)$ 的一个可测映射, 一个 1-维随机变量简称为随机变量. 注意与可测函数不同的是, 随机变量只取有限值. 这时令 $\sigma(\xi):=\xi^{-1}(\mathscr{B})$, 称为是由 ξ 生成的 σ- 代数, 显然它是 Ω 上使得 ξ 成为随机变量的最小 σ- 代数, 类似地如果 $\{\xi_i:i\in I\}$ 是一族随机变量, 那么 Ω 上使得它们都可测的最小 σ- 代数称为是由它们生成的 σ- 代数, 记为 $\sigma(\xi_i:i\in I)$. 实际上, $\sigma(\xi_i:i\in I)$ 等同于子集类 $\bigcup_{i\in I}\sigma(\xi_i)$ 所生成的 σ- 代数, 如果令

$$\mathscr{A}_0:=\{A_1\cap\cdots\cap A_n:n\geqslant 1,A_k\in\sigma(\xi_{i_k}),i_k\in I,1\leqslant k\leqslant n\},$$

那么 $\mathscr{A}_0$ 是 π- 类, 且 $\sigma(\xi_i:i\in I)=\sigma(\mathscr{A}_0)$.

设 ξ 是 Ω 上的 n- 维随机向量, 记 μ_ξ (或记 $\xi(\mathbb{P}),\mathbb{P}\circ\xi^{-1}$) 是概率 $\mathbb{P}$ 在 ξ 下的像测度, 它是 $\mathbf{R}^n$ 上的一个分布, 称为 ξ 的(联合)分布, 对应的分布函数称为是 ξ 的分布函数. 给定分布 μ, 如果存在概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 及其上随机变量 ξ, 使得 $\mu_\xi=\mu$, 称 ξ 是 μ (在概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 上) 的一个实现. 显然任何随机向量都可以实现.

下面的定理说明任何 $\mathbf{R}$ 上的所有分布可以在同一个概率空间上实现.

定理 1.3.2 存在概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$, 使得对 $\mathbf{R}$ 上任何分布函数 F 存在随机变量 ξ 使得 ξ 的分布函数恰是 F.

证明 设 $(\Omega, \mathscr{F}, \mathbb{P})$ 是 $(0,1)$ 上的 Lebesgue 测度构成的概率空间. 对 $\omega \in \Omega$ 定义

$$\xi(\omega) := \inf\{x \in \mathbf{R}: \ F(x) \geqslant \omega\}.$$

因为 F 右连续, 故右侧集合对递减列极限封闭, 因此下确界可以达到. 那么, 对任何 $x \in \mathbf{R}$, $\xi(\omega) \leqslant x$ 等价于 $\omega \leqslant F(x)$. 推出 ξ 是 Ω 上随机变量且其分布函数是 $F(x)$. □

注 上面定义的函数是 F 的(某种意义下)反函数. ξ 是左连续的并且在 ω 点右连续当且仅当 F 在点 $\xi(\omega)$ 的某一右侧邻域上严格递增. (这是一个有趣的数学分析习题.)

两个随机变量称为是同分布的, 如果它们有相同的分布或分布函数. 一个分布通常可以有不同的实现. 不仅是指实现为相同概率空间上的不同随机变量, 而且也可实现在完全不同的概率空间上. 因此在许多情况下, 我们更关心分布, 而不在意它是怎样实现的. 下面给出概率论中一些重要的分布. 在这里分布通常是分类型的, 一个随机变量的分布是某种类型的, 我们说随机变量服从这类型的分布.

例 1.3.3 (单点分布) 取定 $a \in \mathbf{R}$, 恒等于常数 a 的随机变量的分布函数是

$$F(x) = \begin{cases} 0, & x < a; \\ 1, & x \geqslant a; \end{cases}$$

F 是分布函数, 对应的分布是 a 点的单点测度 ϵ_a. ▮

例 1.3.4 设 $0 < p < 1$, 设随机变量取值 1 的概率是 p, 取值 0 的概率是 $1-p$, 那么其分布函数为

$$F(x) = \begin{cases} 0, & x < 0; \\ 1-p, & 0 \leqslant x < 1; \\ 1, & x \geqslant 1; \end{cases}$$

对应的分布是 $(1-p)\epsilon_0 + p\epsilon_1$. 此分布称为是 Bernoulli 分布. ▮

例 1.3.5　说随机变量 X 服从参数为 $\lambda > 0$ 的 Poisson 分布, 如果其分布律为

$$\mathbb{P}(X = k) = \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!},\ k = 0, 1, 2, \cdots.$$

只需验证右边是一个分布律就可以了. Poisson 分布描述某段时间内某事件发生的次数. ∎

例 1.3.6　说随机变量 X 服从区间 (a, b) 上的均匀分布, 如果 X 的分布函数为

$$\mathbb{P}(X \leqslant x) = \begin{cases} 0, & x \leqslant a; \\ \dfrac{x-a}{b-a}, & x \in (a, b); \\ 1, & x \geqslant b. \end{cases}$$

这是一个连续的分布函数. 容易验证, X 落在 (a, b) 内的任何等长度区间上的概率是一样的, 这相当于古典概率的等可能性, 所以称为均匀分布. 是最重要的概率分布之一. ∎

一个 n- 维分布函数 F 称为连续型的, 如果对应的分布关于 $\mathbf{R}^n$ 上的 Lebesgue 测度绝对连续, 即存在 $\mathbf{R}^n$ 上的非负可测函数 f, 使得

$$F(x_1, \cdots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(t_1, \cdots, t_n)\mathrm{d}t_1 \cdots \mathrm{d}t_n,$$

$(x_1, \cdots, x_n) \in \mathbf{R}^n$. 这时称 f 是 F 的密度函数或 f 是一个概率密度, 显然均匀分布函数是连续型的, 而 Bernoulli, Poisson 分布函数不是连续型的.

例 1.3.7　说随机变量 X 服从参数为 $\alpha > 0$ 的指数分布, 如果它有密度函数

$$f(x) = \begin{cases} \alpha\mathrm{e}^{-\alpha x}, & x > 0; \\ 0, & x \leqslant 0. \end{cases}$$

容易验证这个函数是一个密度函数. 通常认为寿命, 如元器件的寿命, 服从指数分布. 指数分布的许多性质类似于离散的几何分布, 如遗忘性. ∎

例 1.3.8　说随机变量 X 服从参数为 a, σ^2 的正态或 Gauss 分布, 如果它有密度函数 (例 1.2.4 说明它的确是密度函数)

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{|x-a|^2}{2\sigma^2}}.$$

记为 $X \sim \mathrm{N}(a,\sigma^2)$, 而 $\mathrm{N}(0,1)$ 对应的分布称为标准正态分布. 我们用 $\varPhi$ 表示标准正态分布的分布函数, 即

$$\varPhi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t.$$

我们知道它不是一个初等函数, 它的值要通过近似计算来获得. 因为其密度函数是偶函数, 因此 $\varPhi(-x) = 1 - \varPhi(x)$, $\varPhi(0) = \frac{1}{2}$. 如果 $X \sim \mathrm{N}(a,\sigma^2)$, 那么容易验证 $\frac{X-a}{\sigma} \sim \mathrm{N}(0,1)$, 因此 X 的分布函数可用 $\varPhi$ 表示:

$$\mathbb{P}(X \leqslant x) = \mathbb{P}\left(\frac{X-a}{\sigma} \leqslant \frac{x-a}{\sigma}\right) = \varPhi\left(\frac{x-a}{\sigma}\right).$$

例 1.3.9 所谓 Cauchy 分布的密度函数为

$$f(x) = \frac{1}{\pi(1+x^2)}, \ -\infty < x < +\infty.$$

例 1.3.10 参数为 $\alpha > 0$ 与 $\beta > 0$ 的 Gamma 分布的密度函数为

$$f(x) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1}\mathrm{e}^{-\beta x}, & x > 0, \\ 0, & x \leqslant 0, \end{cases}$$

其中 $\Gamma(\alpha)$ 是 Gamma 函数

$$\Gamma(\alpha) := \int_0^{+\infty} x^{\alpha-1}\mathrm{e}^{-x}\mathrm{d}x.$$

例 1.3.11 设 $a \in \mathbf{R}^n$ (作为行向量), B 是 n- 阶对称正定矩阵, 定义

$$p(x) := \frac{1}{(2\pi)^{\frac{n}{2}}|B|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(x-a)B^{-1}(x-a)^{\mathrm{T}}\right), \quad x \in \mathbf{R}^n,$$

其中 T 表示转置, $|B|$ 表示 B 的行列式. 容易验证 p 是一个分布函数的密度函数. 我们称由 p 决定的分布函数是 n- 维正态(Guass)分布, 当 $a = 0$ 时, 称为是中心化的正态分布; 当 $a = 0$ 且 B 是单位矩阵时称为是 n- 维标准正态分布.

如果一个随机变量 ξ 关于概率测度 $\mathbb{P}$ 可积, 则其积分 $\mathbb{P}(\xi)$ 通常称为是 ξ 的数学期望或均值, 常理解为 ξ 在 Ω 上的平均, 记为 $\mathbb{E}\xi$. 另外 ξ 在 $A \in \mathscr{F}$ 上的积分也

常记为 $\mathbb{E}(\xi;A)$. 由变量替换公式得

$$\mathbb{E}\xi=\int_{\mathbf{R}}x\mu_\xi(\mathrm{d}x).$$

注意符号 $\mathbb{P}$ 与 $\mathbb{E}$ 没有本质区别, $\mathbb{P}$ 习惯用于事件的概率, 而 $\mathbb{E}$ 用于随机变量的期望, 或者 $\mathbb{P}(A)=\mathbb{E}1_A$. 进一步地, 设 f 是 $\mathbf{R}$ 上 Borel-可测函数, 则 $f\circ\xi$ 也是随机变量, 如果可积, 由变量替换公式, 定理 1.2.2, 有

$$\mathbb{E}f(\xi)=\int_{\mathbf{R}}f(x)\mu_\xi(\mathrm{d}x)=\int_{-\infty}^{+\infty}f(x)\mathrm{d}F_\xi(x).$$

上式右边是 Lebesgue-Stieltjes 意义的积分, 当 f 连续时是 Riemann-Stieltjes 意义的. 随机变量的另一个重要的数字特征是方差. 如果随机变量 ξ 是平方可积的, 即 $\mathbb{E}\xi^2<\infty$, 则 ξ 的方差定义为 $D\xi:=\mathbb{E}(\xi-\mathbb{E}\xi)^2=\mathbb{E}\xi^2-(\mathbb{E}\xi)^2$, 它用来测量随机变量与其数学期望之间的平均偏差.

如果 ξ 是离散的, 那么 $\mathbb{E}\xi=\sum\limits_{x\in R(\xi)}x\mathbb{P}(\xi=x)$. 如果 ξ 是连续型的且密度函数可选为是分段连续的, 那么自然地, 上面的 Riemann-Stieltjes 积分可转化为通常的 Riemann 积分

$$\mathbb{E}f(\xi)=\int_{-\infty}^{+\infty}f(x)F'(x)\mathrm{d}x.$$

例 1.3.12 设 ξ 是 Ω 上服从参数为 λ 的 Poisson 分布的随机变量, 则 ξ 的分布为

$$\mu_\xi=\mathrm{e}^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\epsilon_n,$$

其中 ϵ_n 是点 n 的 Dirac 测度. 容易计算 $\mathbb{E}\xi=\mathrm{e}^{-\lambda}\sum_{n=0}^{\infty}\dfrac{\lambda^n}{n!}n=\lambda$. 同理可得 $\mathbb{E}\xi^2=\lambda+\lambda^2$, 故 $D\xi=\lambda$. ▮

例 1.3.13 设随机变量 ξ 服从 $[a,b]$ 上均匀分布, 其密度函数为 $\dfrac{1}{b-a}1_{[a,b]}$, 因此

$$\mathbb{E}\xi=\frac{1}{b-a}\int_a^b x\mathrm{d}x=\frac{b+a}{2},$$

$$D\xi=\mathbb{E}\xi^2-(\mathbb{E}\xi)^2=\frac{(b-a)^2}{12}.$$

▮

例 1.3.14 (参见例 1.3.8) 设 ξ 是服从参数为 a, σ^2 的正态分布的随机变量, 容易计算 $\mathbb{E}\xi = a$, $D\xi = \sigma^2$. 首先如果 ξ 是标准正态分布的, 那么密度函数是偶函数, 故 $\mathbb{E}\xi = 0$, 由分部积分

$$D\xi = \mathbb{E}\xi^2 = \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} x^2 \mathrm{e}^{-\frac{x^2}{2}} \mathrm{d}x = 1.$$

一般地, 因为 $\dfrac{\xi - a}{\sigma}$ 是标准正态的, 由期望或积分的性质得 $\mathbb{E}\xi = a$, $D\xi = \sigma^2$. ▮

设 ξ, η 是两个平方可积随机变量, 定义

$$\operatorname{cov}(\xi, \eta) := \mathbb{E}(\xi - \mathbb{E}\xi)(\eta - \mathbb{E}\eta),$$

称为是 ξ, η 的协方差. 由 Cauchy-Schwarz 不等式, $|\operatorname{cov}(\xi, \eta)|^2 \leqslant D\xi \cdot D\eta$. 定义 ξ 与 η 的相关系数为

$$\rho(\xi, \eta) := \frac{\operatorname{cov}(\xi, \eta)}{\sqrt{D\xi \cdot D\eta}}.$$

当 $\rho(\xi, \eta) = 0$ 即 $\mathbb{E}\xi\eta = \mathbb{E}\xi\mathbb{E}\eta$ 时, 称 ξ 与 η 不相关. 独立的随机变量一定不相关, 但不相关的随机变量未必独立.

例 1.3.15 设 (ξ, η) 服从 $D = \{x \in \mathbf{R}^2 : |x| \leqslant 1\}$ 上的均匀分布, 即对任何 $A \in \mathscr{B}^2$, 有

$$\mathbb{P}((\xi, \eta) \in A) = \frac{1}{\pi}|A \cap D|,$$

其中 $|\cdot|$ 是 $\mathbf{R}^2$ 上 Lebesgue 测度, 那么 ξ, η 不相关, 但也不独立. ▮

设 $(\xi_1, \cdots, \xi_d)$ 是平方可积的 d- 维随机变量, 令 $c_{ij} := \operatorname{cov}(\xi_i, \xi_j)$, $1 \leqslant i, j \leqslant d$. 那么 (c_{ij}) 称为是这个随机向量的协方差矩阵. 不难验证协方差矩阵一定是非负定的. 协方差矩阵是随机向量的一个重要数字特征.

最后我们介绍独立的概念以及简单而重要的 Borel-Cantelli 引理, 它是概率论中独有的. 设 $(\Omega, \mathscr{F}, \mathbb{P})$ 是概率空间, $\{\mathscr{A}_i : i \in I\}$ 是 $\mathscr{F}$ 中子类的集合. $\{\mathscr{A}_i\}$ 称为是相互独立的, 如果对任何 I 的有限子集 I_0 与任何 $A_i \in \mathscr{A}_i$, $i \in I_0$, 有

$$\mathbb{P}\left(\bigcap_{i \in I_0} A_i\right) = \prod_{i \in I_0} \mathbb{P}(A_i);$$

事件集 $\{A_i\} \subset \mathscr{F}$ 相互独立, 如果作为子类的集合 $\{\{A_i\}\}$ 是相互独立的; 设对任意 $i \in I$, $\mathscr{H}_i$ 是 Ω 上的随机变量集, 称 $\{\mathscr{H}_i\}$ 相互独立, 如果 $\{\sigma(\mathscr{H}_i) : i \in I\}$ 是相互独

立的. 显然两个随机变量 ξ, η 独立当且仅当对任何 $x, y \in \mathbf{R}$,

$$\mathbb{P}(\xi \leqslant x, \eta \leqslant y) = \mathbb{P}(\xi \leqslant x)\mathbb{P}(\eta \leqslant y).$$

定理 1.3.3 设 $\{\mathscr{A}_i : i \in I\}$ 是 $\mathscr{F}$ 中相互独立的子集类, 如果每个 $\mathscr{A}_i$ 都是 π- 类, 那么 $\{\sigma(\mathscr{A}_i) : i \in I\}$ 是相互独立的.

定理的证明应用引理 1.1.1 的 Dynkin 定理立刻可得, 留作习题. 下面是一个简单而重要的结果.

定理 1.3.4 (Borel-Cantelli) 设 $\{A_n\}$ 是事件列.

(1) 若 $\sum\limits_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, 则 $\mathbb{P}(\limsup_n A_n) = 0$;

(2) 若 $\{A_n\}$ 是独立事件列且 $\sum\limits_{n} \mathbb{P}(A_n) = \infty$, 则 $\mathbb{P}(\limsup_n A_n) = 1$.

证明 (1) 首先 $\mathbb{P}(\limsup_n A_n) = \lim_n \mathbb{P}(\bigcup_{k \geqslant n} A_k)$, 而

$$\mathbb{P}(\bigcup_{k \geqslant n} A_k) \leqslant \sum_{k \geqslant n} \mathbb{P}(A_k) \longrightarrow 0,$$

因为级数 $\sum\limits_{n=1}^{\infty} \mathbb{P}(A_n)$ 收敛.

(2) 对 $n < N$, 由于 $\{A_n\}$ 独立,

$$\mathbb{P}(\bigcap_{k=n}^{N} A_k^c) = \prod_{k=n}^{N} (1 - \mathbb{P}(A_k)) \leqslant \prod_{k=n}^{N} \mathrm{e}^{-\mathbb{P}(A_k)} = \mathrm{e}^{-\sum\limits_{k=n}^{N} \mathbb{P}(A_k)}.$$

得 $\lim_N \mathbb{P}(\bigcap_{k=n}^{N} A_k^c) = 0$, 即 $\mathbb{P}(\bigcup_{k=n}^{\infty} A_k) = 1$, 故 $\mathbb{P}(\limsup_n A_n) = 1$. □

将 Fatou 引理应用于 $\{1_{A_n}\}$ 推出, $\liminf_n \mathbb{P}(A_n) = 0$ 蕴含 $\mathbb{P}(\liminf_n A_n) = 0$. 这是一个条件与结论都弱于 Borel-Cantelli 引理的类似结果.

习 题

1. (1) 在一条线段 AB 上任取 3 点 X, Y, Z, 求线段 AX, AY, AZ 能够组成三角形的概率.

   (2) 在一个线段上任取两点自然形成 3 条线段, 求它们能组成三角形的概率.

   (3) 一根木棍随机断为两段, 然后较长的一根又随机地断为两段, 求此三段能组成三角形的概率.

2. 如果随机变量 ξ 有一个连续的分布函数 F, 那么 $F(\xi)$ 服从 $(0,1)$ 上均匀分布.

3. 设独立随机变量 $X_1,\cdots,X_n$ 服从标准正态分布, 证明 $\frac{1}{\sqrt{n}}(X_1+\cdots+X_n)$ 也服从标准正态分布.

4. 上述问题的反问题: 设 F 是一个分布函数, 使得对任何 $n\geqslant 1$, $X_1,\cdots,X_n$ 独立且分布函数都为 F 时有 $\frac{1}{\sqrt{n}}(X_1+\cdots+X_n)$ 的分布也是 F, 问 F 是什么样的?

5. 设 $X_1,\cdots,X_n$ 是服从 Cauchy 分布的独立随机变量, 证明 $\frac{1}{n}(X_1+\cdots+X_n)$ 也服从 Cauchy 分布. 问类似正态分布的反问题是否对 Cauchy 分布成立?

6. 如果 X 服从参数为 λ 的 Poisson 分布, Y 服从参数为 $\alpha,1$ 的 Gamma 分布, 那么对 $\alpha=1,2,\ldots$, 有 $\mathbb{P}(X\geqslant\alpha)=\mathbb{P}(Y\leqslant\lambda)$.

7. 设独立随机变量 $X_1,\cdots,X_n$ 都服从参数为 α 的指数分布, 证明: $X_1+\cdots+X_n$ 服从 Gamma 分布.

8. 设独立随机变量 $X_1,\cdots,X_n$ 服从标准正态分布, 则

   (1) $X_1^2+\cdots+X_n^2$ 服从 Gamma 分布;

   (2) 设 $\overline{X}:=\frac{1}{n}(X_1+\cdots+X_n)$, $S:=\frac{1}{n-1}[(X_1-\overline{X})^2+\cdots+(X_n-\overline{X})^2]$, 则 $\overline{X}$ 与 S 独立, 分别服从正态分布与 Gamma 分布.

9. 设 $(\Omega,\mathscr{F},\mathbb{P})$ 是一个概率空间. 取 $\Omega_0\in\mathscr{F}$, $\mathbb{P}(\Omega_0)>0$, 定义

$$\mathbb{P}_{|\Omega_0}(A):=\mathbb{P}(A|\Omega_0),\quad A\in\mathscr{F}.$$

   证明: (1) $(\Omega,\mathscr{F},\mathbb{P}_{|\Omega_0})$ 也是一个概率空间, 称为是关于 Ω_0 的条件概率空间.

   (2) 对任何 $B\in\mathscr{F}$, $\mathbb{P}_{|\Omega_0}(A|B)=\mathbb{P}(A|B\cap\Omega_0)$, 当条件概率有意义时.

10. 设 μ 是一个分布, F 是对应的分布函数. 证明: $\mu(\{x\})=F(x)-F(x-)$, $x\in\mathbf{R}$.

11. 举例说明连续的分布函数未必是连续型的.

12. 随机变量 ξ,η 称为不相关, 如果 $\mathbb{E}(\xi\cdot\eta)=\mathbb{E}\xi\cdot\mathbb{E}\eta$. 证明: 两个独立随机变量不相关.

13. 设 ξ 服从单位区间上均匀分布, ξ_n 是 ξ 的二进制表示的第 n 位小数. 将 $\{\xi_n\}$ 重新编号为 $\{\xi_{k,n}:\ k,n\geqslant 1\}$, 令

$$\eta_k := \sum_{n=1}^{\infty} \frac{\xi_{k,n}}{2^n},\ k \geqslant 1.$$

证明: $\{\eta_k : k \geqslant 1\}$ 是均匀分布的独立随机序列.

14. 设 ξ 服从 $[0, 2\pi)$ 上的均匀分布, 令 $X = \cos\xi$, $Y = \sin\xi$. 求 $\rho(X, Y)$ 与 $\rho(X^2, Y^2)$.

15. (1) 在一个圆心为 O 的单位圆周上任取两点 A, B, 求形成的三角形 OAB 面积的分布与期望.

    (2) 在单位圆周上任取 3 点 A, B, C, 求三角形 ABC 面积的期望.

16. 甲袋中有 1 个黑球 2 个白球, 乙袋中有 3 个白球, 每次从两袋中各任取一球, 交换放入另一袋中, 求交换 n 次后, 黑球仍在甲袋中的概率.

17. 设 ξ 是一个服从 $[0,1]$ 上均匀分布的随机变量. 问: 是否能在 $[0,1]$ 上构造函数 f 使得 $f(\xi)$ 服从参数为 λ 的 Poisson 分布?

18. 证明: (1) 一个概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 称为非原子的, 若对任何 $A \in \mathscr{F}$, $\mathbb{P}(A) > 0$, 存在 $B \in \mathscr{F}$, $B \subset A$ 使得 $\mathbb{P}(A) > \mathbb{P}(B) > 0$. 证明: 这时 $\mathbb{P}$ 的值域是 $[0,1]$.

    (2) 概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 的概率 $\mathbb{P}$ 的值域总是闭的.

19. 设 μ, ν 是两个 d- 维分布, 定义卷积 $*$:

$$\mu * \nu(A) := \int 1_A(x+y)\mathrm{d}\mu \times \nu = \int_{x+y\in A} \mu(\mathrm{d}x)\nu(\mathrm{d}y),$$

    其中 $A \in \mathscr{B}^d$. 证明:

    (1) $\mu * \nu$ 是一个 d- 维分布;

    (2) 设 ξ, η 是相同概率空间上的独立随机变量, 分布分别为 μ, ν, 则 $\xi + \eta$ 的分布为 $\mu * \nu$. 反之不对, 举例.

20. 设 $(\Omega, \mathscr{F}, \mathbb{P})$ 是概率空间, $\mathscr{A}_i \subset \mathscr{F}$, $i \in I$ 是 π 类, 则 $\{\mathscr{A}_i : i \in I\}$ 独立当且仅当 $\{\sigma(\mathscr{A}_i) : i \in I\}$ 独立.

21. 随机变量集 $\{\xi_1, \cdots, \xi_n\}$ 与 $\{\eta_1, \cdots, \eta_m\}$ 独立当且仅当对任意非负可测函数 $f \in \mathscr{B}^n$, $g \in \mathscr{B}^m$, 有

$$\begin{aligned}&\mathbb{E}[f(\xi_1, \cdots, \xi_n)g(\eta_1, \cdots, \eta_m)]\\&\qquad = \mathbb{E}f(\xi_1, \cdots, \xi_n)\mathbb{E}g(\eta_1, \cdots, \eta_m).\end{aligned}$$

22. 随机变量 ξ, η 独立当且仅当对任何 $\mathbf{R}$ 上有界连续函数 f, g, 有

$$\mathbb{E}[f(\xi)g(\eta)] = \mathbb{E}f(\xi)\mathbb{E}g(\eta).$$

23. 验证随机向量的协方差矩阵一定是对称非负定的.
24. 设 $(\xi_1, \cdots, \xi_n)$ 服从例 1.3.11 中的 Gauss 分布, 证明: 它的协方差矩阵恰好是 B.
25. 设 ξ, η 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 两平方可积随机变量, 满足对任何 $w_1, w_2 \in \Omega$, 有 $(\xi(w_1) - \xi(w_2))(\eta(w_1) - \eta(w_2)) \geqslant 0$. 证明: $\mathrm{cov}(\xi, \eta) \geqslant 0$.
26. 设 $\{A_n\}$ 是独立事件列, $\xi(\omega) := \#\{n : \omega \in A_n\}$. 证明: $\mathbb{P}(\xi < \infty) > 0$ 蕴含 $\mathbb{E}\xi < \infty$.
27. (Kolmogorov 0-1 律)设 $\xi_1, \cdots, \xi_n, \cdots$ 是独立随机变量序列, 令

$$\mathscr{F} := \sigma(\xi_1, \xi_2, \cdots), \ \mathscr{A} := \bigcap_n \sigma(\xi_n, \xi_{n+1}, \cdots).$$

证明: $\mathscr{F}$ 与 $\mathscr{A}$ 独立, 且对任何 $A \in \mathscr{A}$, $\mathbb{P}(A) = 0$ 或 1.

28. 设有概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$, 对给定的 $0 < a < \dfrac{1}{2}$, 存在独立事件列 $\{A_n\} \in \mathscr{F}$ 满足 $a < \mathbb{P}(A_n) < 1 - a$, 证明: 尾 σ- 代数 $\bigcap_n \sigma(\{A_n, A_{n+1}, \cdots\})$ 不是可列生成的.

§1.4 随机变量的收敛性

对 $r \geqslant 1$, 用 $L^r(\Omega, \mathscr{F}, \mathbb{P})$ (简写为 $L^r(\Omega)$) 表示具有有限 r- 阶绝对矩的随机变量全体, 即

$$L^r(\Omega) := \{\xi : \ \xi \text{ 是随机变量且 } \mathbb{E}|\xi|^r < \infty\}.$$

对于 $\xi \in L^r(\Omega)$, 令 $\|\xi\|_r := (\mathbb{E}|\xi|^r)^{\frac{1}{r}}$, 则 $L^r(\Omega)$ 关于范数 $\|\cdot\|_r$ 是一个 Banach 空间 (注意在此我们将不区别几乎处处相等的随机变量). 我们引用 Banach 空间中强弱拓扑的概念. 下面是重要的 Hölder 不等式.

定理 1.4.1 (Hölder) 设 $1 < r < \infty$, $1 < s < \infty$ 且 $\dfrac{1}{r} + \dfrac{1}{s} = 1$, 若 $\xi \in L^r(\Omega)$, $\eta \in L^s(\Omega)$, 则 $\xi\eta \in L^1(\Omega)$ 且 $\|\xi\eta\|_1 \leqslant \|\xi\|_r \cdot \|\eta\|_s$.

证明 利用指数函数的凸性,

$$e^{x+y} \leqslant \frac{1}{r}e^{rx} + \frac{1}{s}e^{sy},\ x, y \in \mathbf{R},$$

首先对任何实数 a, b, 有

$$|ab| \leqslant \frac{1}{r}|a|^r + \frac{1}{s}|b|^s.$$

令 $a = \dfrac{\xi}{\|\xi\|_r}$, $b = \dfrac{\eta}{\|\eta\|_s}$ 我们得到

$$\frac{|\xi\eta|}{\|\xi\|_r\|\eta\|_s} \leqslant \frac{1}{r}\frac{|\xi|^r}{\|\xi\|_r^r} + \frac{1}{s}\frac{|\eta|^s}{\|\eta\|_s^s}.$$

由此得

$$\mathbb{E}\frac{|\xi\eta|}{\|\xi\|_r \cdot \|\eta\|_s} \leqslant 1,$$

即是 Hölder 不等式. □

当 $r = 2$ 时, $(\mathbb{E}|\xi\eta|)^2 \leqslant \mathbb{E}|\xi|^2 \cdot \mathbb{E}|\eta|^2$, 该不等式也称为 Cauchy-Schwarz 不等式. 另外还有一个简单但重要的 Chebyshev 不等式.

定理 1.4.2 (Chebyshev)　设 ξ 是一个随机变量, $r > 0$, 则对于任意 $\epsilon > 0$, 有

$$\mathbb{P}(\{|\xi| \geqslant \epsilon\}) \leqslant \frac{1}{\epsilon^r}\mathbb{E}|\xi|^r.$$

证明　由定义

$$\mathbb{P}(\{|\xi| \geqslant \epsilon\}) \leqslant \mathbb{E}\left[\left(\frac{|\xi|}{\epsilon}\right)^r; \{|\xi| \geqslant \epsilon\}\right] \leqslant \frac{1}{\epsilon^r}\mathbb{E}|\xi|^r.$$

完成证明. □

Chebyshev 不等式的一个简单推论: $\mathbb{E}|X| < +\infty$ 蕴含着 $|X| < +\infty$ a.s..

定义 1.4.1　设 $\{\xi_n\}$ 是一个随机变量序列, ξ 是一个随机变量.

(1) 称 $\{\xi_n\}$ 依概率收敛于 ξ, 如果对任何 $\epsilon > 0$,

$$\lim_n \mathbb{P}(\{\omega \in \Omega : |\xi_n(\omega) - \xi(\omega)| \geqslant \epsilon\}) = 0.$$

记为 $\xi_n \xrightarrow{p} \xi$.

(2) 称 $\{\xi_n\}$ 几乎处处(或概率 1)收敛于 ξ, 如果

$$\mathbb{P}(\{\omega \in \Omega : \lim_n \xi_n(\omega) = \xi(\omega)\}) = 1.$$

记为 $\xi_n \xrightarrow{\text{a.s.}} \xi$.

(3) 称 $\{\xi_n\}$ L^r- 收敛于 ξ $(r \geqslant 1)$, 如果 $\xi_n, \xi \in L^r(\Omega)$ 且

$$\lim_n \mathbb{E}|\xi_n - \xi|^r = 0,$$

或者说, $\{\xi_n\}$ 在 $L^r(\Omega)$ 中范收敛于 ξ, 记为 $\xi_n \xrightarrow{L^r} \xi$.

我们来分析这几种收敛相互之间的关系. 首先不难看出这几种收敛的极限在几乎处处相等的意义之下是唯一的. 由 Chebyshev 不等式容易看出, 如果 $\xi_n \xrightarrow{L^r} \xi$ (对某个 $r > 0$), 则 $\xi_n \xrightarrow{p} \xi$. 为了弄清依概率收敛与几乎处处收敛之间的关系, 首先容易验证

$$\{\omega : \lim \xi(\omega) = \xi(\omega)\} = \bigcap_{\epsilon > 0} \bigcup_{N \geqslant 1} \bigcap_{n \geqslant N} \{|\xi_n - \xi| < \epsilon\}.$$

因此如果 $\xi_n \xrightarrow{\text{a.s.}} \xi$, 则对任意 $\epsilon > 0$, 有

$$\mathbb{P}(\bigcup_{N \geqslant 1} \bigcap_{n \geqslant N} \{|\xi_n - \xi| < \epsilon\}) = 1.$$

从而由 Fatou 引理

$$\liminf_n \mathbb{P}(\{|\xi_n - \xi| < \epsilon\}) \geqslant \mathbb{P}(\liminf_n \{|\xi_n - \xi| < \epsilon\}) = 1.$$

故 $\lim_n \mathbb{P}(\{|\xi_n - \xi| < \epsilon\}) = 1$, 即 $\xi_n \xrightarrow{p} \xi$, 几乎处处收敛蕴含依概率收敛.

定理 1.4.3 设 $\{\xi_n\}$ 是一个随机变量序列, ξ 是一个随机变量.

(1) $\xi_n \xrightarrow{L^r} \xi$ (对某个 $r > 0$) 蕴含着 $\xi_n \xrightarrow{p} \xi$;

(2) $\xi_n \xrightarrow{\text{a.s.}} \xi$ 蕴含着 $\xi_n \xrightarrow{p} \xi$;

(3) 若 $\xi_n \xrightarrow{p} \xi$, 则存在 $\{\xi_n\}$ 的一个子序列 $\{\xi_{n_k}\}$ 几乎处处收敛于 ξ.

证明 只需证 (3). 对任何整数 $k > 0$, 必存在 n_k 使得

$$\mathbb{P}\left(\left\{|\xi_{n_k} - \xi| > \frac{1}{k}\right\}\right) \leqslant \frac{1}{2^k}.$$

由 Borel-Cantelli 引理, 集 $N := \bigcap_{K \geqslant 1} \bigcup_{k \geqslant K} \{|\xi_{n_k} - \xi| > \dfrac{1}{k}\}$, 概率为 0. 而显然 $N^c \subset \{\lim_k \xi_{n_k} = \xi\}$. 因此 $\xi_{n_k} \xrightarrow{\text{a.s.}} \xi$. □

例 1.4.1 设 $\Omega = [0, 1]$, $\mathscr{F}$ 是其上的 Borel 域, $\mathbb{P}$ 是 Lebesgue 测度. 记

$$E_k^m := \left[\frac{k-1}{2^m}, \frac{k}{2^m}\right], \quad m \geqslant 0,\ 1 \leqslant k \leqslant 2^m.$$

再令 $\xi_1 := 1_{E_1^0}$, $\xi_2 := 1_{E_1^1}$, $\xi_3 := 1_{E_2^1}, \cdots$, 一般地, 如果 $2^m \leqslant n < 2^{m+1}$, 令

$$\xi_n := 1_{E^m_{n-2^m+1}}.$$

当然 $\{\xi_n\}$ 是非负随机变量序列, 容易看出 $\mathbb{E}(|\xi_n|^r) = \dfrac{1}{2^m}$, 故 $\xi_n \xrightarrow{L^r} 0$ (任意 $r \geqslant 1$), 但是 $\{\xi_n\}$ 不是几乎处处收敛的. ▮

例 1.4.2　设概率空间如上, $\xi_n := n1_{[0,\frac{1}{n}]}$. 则容易验证 $\xi_n \xrightarrow{\text{a.s.}} 0$, 但是 $\{\xi_n\}$ 不是 L^r-收敛的. ▮

独立地重复一个成功概率为 p 的 Bernoulli 试验, 用 X_n 表示前 n 次试验时成功的次数, 那么 $\dfrac{X_n}{n}$ 是前 n 次试验成功的频率, 也是一个随机变量, 下面定理说明这个频率在某种意义下收敛于成功的概率 p, 即大数定律. 这是 Bernoulli 在 1713 年发表的, 据称是概率研究的第一篇论文.

定理 1.4.4 (Bernoulli)　对任何 $\epsilon > 0$, 有

$$\lim_n \mathbb{P}\left(\left|\frac{X_n}{n} - p\right| > \epsilon\right) = 0.$$

定理用 Chebyshev 不等式容易验证. 一般地, 设 $\{\xi_n\}$ 是可积随机变量序列. 如果 $\dfrac{1}{n}\sum\limits_{i=1}^n(\xi_i - \mathbb{E}\xi_i) \xrightarrow{p} 0$, 称 $\{\xi_n\}$ 满足大数定律. 如果 $\dfrac{1}{n}\sum\limits_{i=1}^n(\xi_i - \mathbb{E}\xi_i) \xrightarrow{\text{a.s.}} 0$, 称 $\{\xi_n\}$ 满足强大数定律. 显然, Chebyshev 不等式可以推出如果 $\{\xi_n\}$ 是独立同分布的平方可积随机序列, 那么它满足大数定律. 下面的结果是 E. Borel 在 1909 年发表的.

定理 1.4.5 (Borel)　如果 $\{\xi_n\}$ 是独立同分布的平方可积随机序列且 $\mathbb{E}\xi_1^4 < \infty$, 那么 $\{\xi_n\}$ 满足强大数定律.

证明　我们下面利用 Chebyshev 不等式和 Borel-Cantelli 引理证明

$$\mathbb{P}\left(\bigcup_{\epsilon>0}\bigcap_{n\geqslant 1}\bigcup_{k\geqslant n}\left\{\left|\frac{1}{k}\sum_{i=1}^k \xi_i - \mathbb{E}\xi_1\right| > \epsilon\right\}\right) = 0.$$

也就是要证明对任何 $\epsilon > 0$,

$$\mathbb{P}\left(\limsup_n \left\{\left|\frac{1}{n}\sum_{i=1}^n \xi_i - \mathbb{E}\xi_1\right| > \epsilon\right\}\right) = 0.$$

而由 Chebyshev 不等式,

$$\begin{aligned}
\sum_n \mathbb{P}\left(\left\{\left|\frac{1}{n}\sum_{i=1}^n \xi_i - \mathbb{E}\xi_1\right| > \epsilon\right\}\right) &\leqslant \sum_n \frac{1}{(\epsilon n)^4}\mathbb{E}\left(\sum_{i=1}^n (\xi_i - \mathbb{E}\xi_i)\right)^4 \\
&= \sum_n \frac{1}{(\epsilon n)^4}\left(\mathbb{E}\sum_{i=1}^n (\xi_i - \mathbb{E}\xi_i)^4 + \sum_{1\leqslant i\neq j\leqslant n} \mathbb{E}(\xi_i - \mathbb{E}\xi_i)^2(\xi_j - \mathbb{E}\xi_j)^2\right) \\
&\leqslant \sum_n \frac{1}{(\epsilon n)^4}(n\mathbb{E}(\xi_1 - \mathbb{E}\xi_1)^4 + n(n-1)(D\xi_1)^2) < \infty.
\end{aligned}$$

因此结论由 Borel-Cantelli 引理推出. □

为了进一步阐述收敛之间的关系, 我们引入一致可积的概念, 它本身也是概率论中最为重要的概念之一.

定义 1.4.2 一个可积随机变量族 $\{\xi_i : i \in I\}$ 称为是一致可积的, 若

$$\lim_{N\to\infty} \sup_I \mathbb{E}(|\xi_i|; |\xi_i| \geqslant N) = 0.$$

显然, 如若 $\{\xi_i : i \in I\}$ 被一个可积随机变量所控制, 则 $\{\xi_i\}$ 是一致可积的. 下面定理给出一致可积的一个等价条件.

定理 1.4.6 设 $\{\xi_i : i \in I\}$ 是可积随机变量族, 则它是一致可积的充要条件是

(1) 一致绝对连续: 对任何 $\epsilon > 0$, 存在 $\delta > 0$ 使得当 $A \in \mathscr{F}$, $\mathbb{P}(A) < \delta$ 时, $\sup_{i\in I} \mathbb{E}(|\xi_i|; A) < \epsilon$;

(2) L^1- 有界: $\sup_{i\in I} \mathbb{E}|\xi_i| < \infty$.

证明 必要性. 对任意 $A \in \mathscr{F}$, $N > 0$, 有

$$\begin{aligned}
\mathbb{E}(|\xi_i|; A) &= \mathbb{E}(|\xi_i|; A \cap \{|\xi_i| \geqslant N\}) + \mathbb{E}(|\xi_i|; A \cap \{|\xi_i| < N\}) \\
&\leqslant \mathbb{E}(|\xi_i|; \{|\xi_i| \geqslant N\}) + N\mathbb{P}(A).
\end{aligned}$$

运用一致可积性, 推出 $\{\xi_i\}$ 是一致绝对连续的. 再在上式令 $A = \Omega$, 得

$$\mathbb{E}|\xi_i| \leqslant \mathbb{E}(|\xi_i|; \{|\xi_i| \geqslant N\}) + N,$$

得到 $\{\xi_i\}$ 的 L^1- 有界性.

充分性. 设 $\{\xi_i\}$ 是一致绝对连续且 L^1- 有界的. 由 Chebyshev 不等式, 当 $N \to \infty$ 时,

$$\sup_i \mathbb{P}(|\xi_i| \geqslant N) \leqslant \frac{1}{N} \sup_i \mathbb{E}|\xi_i| \longrightarrow 0.$$

从而由 $\{\xi_i\}$ 的一致绝对连续性得到, 对任何 $\epsilon > 0$, 存在 $N > 0$, 使得 $\mathbb{E}(|\xi_i|; \{|\xi_i| \geqslant N\}) \leqslant \epsilon$, 即一致可积性. □

定理 1.4.7 可积随机变量序列 $\{\xi_n\}$ L^1- 收敛于 ξ 的充要条件是 $\{\xi_n\}$ 是一致可积的且 $\xi_n \xrightarrow{p} \xi$.

证明 必要性. 首先 $\xi_n \xrightarrow{p} \xi$ 是显然的, $\{\xi_n\}$ 的 L^1- 有界性也是显然的. 而 $\{\xi_n\}$ 的一致绝对连续性由下面的不等式及 ξ 是可积的事实立即推出. 对任意 $A \in \mathscr{F}$,

$$\mathbb{E}(|\xi_n|; A) \leqslant \mathbb{E}(|\xi|; A) + \mathbb{E}(|\xi_n - \xi|).$$

充分性. 对任意 $\epsilon > 0$,

$$\begin{aligned}
\mathbb{E}|\xi_n - \xi| &\leqslant \mathbb{E}(|\xi_n - \xi|; \{|\xi_n - \xi| < \epsilon\}) + \mathbb{E}(|\xi_n - \xi|; \{|\xi_n - \xi| \geqslant \epsilon\}) \\
&\leqslant \epsilon + \mathbb{E}(|\xi_n|; \{|\xi_n - \xi| \geqslant \epsilon\}) + \mathbb{E}(|\xi|; \{|\xi_n - \xi| \geqslant \epsilon\}).
\end{aligned}$$

因为 $\lim_n \mathbb{P}(\{|\xi_n - \xi| \geqslant \epsilon\}) = 0$, 故由 $\{\xi_n\}$ 的一致绝对连续性和 ξ 的绝对连续性推出右边可以任意地小, 因此 $\xi_n \xrightarrow{L^1} \xi$. □

下面我们讨论分布收敛的问题.

定理 1.4.8 设 $\{\xi_n\}$, ξ 是随机变量, 分别具有分布函数 $\{F_n\}$, F. 如果 $\xi_n \xrightarrow{p} \xi$, 则对 F 的任意连续点 x, 有 $\lim_n F_n(x) = F(x)$.

证明 对任意 $\epsilon > 0$,

$$\begin{aligned}
F_n(x) = \mathbb{P}(\xi_n \leqslant x) &= \mathbb{P}(\{\xi_n \leqslant x\} \cap \{|\xi_n - \xi| < \epsilon\}) + \mathbb{P}(\{\xi_n \leqslant x\} \cap \{|\xi_n - \xi| \geqslant \epsilon\}) \\
&\leqslant \mathbb{P}(\xi \leqslant x + \epsilon) + \mathbb{P}(|\xi_n - \xi| \geqslant \epsilon) \\
&= F(x + \epsilon) + \mathbb{P}(|\xi_n - \xi| \geqslant \epsilon);
\end{aligned}$$

$$F_n(x) = \mathbb{P}(\xi_n \leqslant x) \geqslant \mathbb{P}(\{\xi \leqslant x - \epsilon\}) \cap \{|\xi_n - \xi| < \epsilon\})$$

$$\geqslant F(x-\epsilon) - \mathbb{P}(|\xi_n - \xi| \geqslant \epsilon).$$

因此

$$F(x-\epsilon) \leqslant \liminf_n F_n(x) \leqslant \limsup_n F_n(x) \leqslant F(x+\epsilon).$$

当 F 在 x 点连续时, 显然有 $\lim_n F_n(x) = F(x)$. □

由以上定理, 我们来定义另一种收敛性.

定义 1.4.3 (1) 设 F_n 是 $\mathbf{R}$ 上分布函数. 称 $\{F_n\}$ 弱收敛, 如果存在一个递增右连续函数 F, 使得对任何 F 的连续点 x, 有 $\lim_n F_n(x) = F(x)$, 记 $F_n \xrightarrow{w} F$. 如果 F 是一个分布函数, 称 F_n 正常地弱收敛(于 F).

(2) 说一个随机变量序列 $\{\xi_n\}$ 依分布收敛于随机变量 ξ, 记 $\xi_n \Rightarrow \xi$, 如果 ξ_n 的分布函数序列 F_n 弱收敛于 ξ 的分布函数 F.

由 Fatou 引理, 分布函数列弱收敛的极限的总变差不会大于 1. 定理 1.4.8 说明依测度收敛蕴含着依分布收敛, 但反之不对, 因为不同随机变量可以有相同分布. 下面的 Skorohod 定理说明弱收敛的分布函数列可以实现为同一个概率空间上的处处收敛的随机变量序列.

定理 1.4.9 (Skorohod) 设 $F_n, n \geqslant 1, F$ 是 $\mathbf{R}$ 上分布函数. 如果 $F_n \xrightarrow{w} F$, 则存在概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 与其上的随机变量 $\{\xi_n\}$, ξ, 使得

(1) 对任何 $\omega \in \Omega$, $\xi_n(\omega) \longrightarrow \xi(\omega)$;

(2) $F_n, n \geqslant 1, F$ 分别是 $\xi_n, n \geqslant 1, \xi$ 的分布函数.

证明 设 $\Omega = (0,1)$, $\mathscr{F} = \mathscr{B}((0,1))$, $\mathbb{P}$ 是其上 Lebesgue 测度. 仿照定理 1.3.2 的证明, 对 $\omega \in \Omega$ 定义

$$\xi(\omega) := \inf\{x \in \mathbf{R}: \ F(x) \geqslant \omega\},$$

$$\xi_n(\omega) := \inf\{x \in \mathbf{R}: \ F_n(x) \geqslant \omega\}, \ n \geqslant 1.$$

则 (2) 成立.

取 ξ 的连续点 ω, 再取 F 的连续点 x, x', 使得 $x < \xi(\omega) < x'$, 由定理 1.3.2 下的注, $F(x) < \omega < F(x')$. 则 n 充分大时, $F_n(x) < \omega < F_n(x')$. 因此 $x < \xi_n(\omega) \leqslant x'$, (想想这里为何是 $\leqslant$ 而不是 $<$?) 故 $|\xi_n(\omega) - \xi(\omega)| \leqslant x' - x$. 推出 $\xi_n(\omega) \to \xi(\omega)$. 因为 ξ 递增, 其不连续点可列, Lebesgue 测度为零. 改变此点集上的值不改变分布或分布函数. 因此 (1) 成立. □

定理 1.4.10　设分布函数 F_n, F 分别对应于分布 μ_n, μ, 则

(1) $F_n \xrightarrow{w} F$ 当且仅当对任何 $f \in C_b(\mathbf{R})$, $\mu_n(f) \to \mu(f)$, 也称 μ_n 弱收敛于 μ, 记 $\mu_n \xrightarrow{w} \mu$;

(2) 如果 $\{f_i : i \in I\}$ 是 $\mathbf{R}$ 上一致有界且等度连续的函数族, 且 $\mu_n \xrightarrow{w} \mu$, 那么收敛 $\mu_n(f_i) \to \mu(f_i)$ 关于 i 是一致的.

证明　(1) 设 $F_n \xrightarrow{w} F$, 则取定理 1.4.9 中的 ξ_n 与 ξ, 对 $f \in C_b(\mathbf{R})$, 应用有界收敛定理 $\mu_n(f) = \mathbb{E}f(\xi_n) \to \mathbb{E}f(\xi) = \mu(f)$. 即 $\mu_n \xrightarrow{w} \mu$. 反过来, 任取 $x_0 < x_1$, 取 $f \in C_b(\mathbf{R})$ 满足 $1_{(-\infty, x_0]} \leqslant f \leqslant 1_{(-\infty, x_1]}$, 由弱收敛性得,

$$F(x_0) \leqslant \mu(f) = \lim_n \mu_n(f) \leqslant \liminf_n F_n(x_1),$$
$$F(x_1) \geqslant \mu(f) = \lim_n \mu_n(f) \geqslant \limsup_n F_n(x_0).$$

即对任何 $x \in \mathbf{R}$, $\delta > 0$, 有

$$F(x-\delta) \leqslant \liminf_n F_n(x) \leqslant \limsup F_n(x) \leqslant F(x+\delta).$$

推出若 F 在 x 连续, 必有 $\lim F_n(x) = F(x)$.

(2) 设 M 是 $\{f_i\}$ 的一致界, 则对任何 $\delta > 0$,

$$\mu_n(f_i) - \mu(f_i)| \leqslant \mathbb{E}(|f_i(\xi_n) - f_i(\xi)|; |\xi_n - \xi| < \delta) + 2M \cdot \mathbb{P}(|\xi_n - \xi| \geqslant \delta).$$

由 $\{f_i\}$ 的等度连续性, 上面的收敛对 i 是一致的. □

例 1.4.3　令

$$p(t,x) := \frac{1}{(2\pi t)^{\frac{d}{2}}} \mathrm{e}^{-\frac{|x|^2}{2t}},\ t > 0, x \in \mathbf{R}^d.$$

对 $t > 0$, 令 $b_t(A) := \int_A p(t,x)\mathrm{d}x$, $A \in \mathscr{B}(\mathbf{R}^d)$ 对 $f \in C_b(\mathbf{R}^d)$, 由控制收敛定理,

$$\begin{aligned}\lim_{t\downarrow 0} b_t(f) &= \lim_{t\downarrow 0} \int_{\mathbf{R}} \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{|x|^2}{2t}} f(x)\mathrm{d}x \\ &= \lim_{t\downarrow 0} \int_{\mathbf{R}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{|x|^2}{2t}} f(\sqrt{t}x)\mathrm{d}x\end{aligned}$$

$$= \int_{\mathbf{R}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{|x|^2}{2t}} \lim_{t\downarrow 0} f(\sqrt{t}x)\mathrm{d}x$$

$$= f(0).$$

而对任何分布 μ, 定义 $f\bar{*}\mu(x) := \int f(x+y)\mu(dy)$, $x \in \mathbf{R}^d$, 那么 $f\bar{*}\mu$ 有界连续, 故

$$b_t * \mu(f) = b_t(f\bar{*}\mu) \to f\bar{*}\mu(0) = \mu(f),$$

即 $b_t * \mu \xrightarrow{w} \mu$.

我们应注意到, 一个分布函数列可以处处收敛于非分布函数. 如 $F_n = 1_{[n,\infty)}$, 则 $F_n \xrightarrow{w} 0$.

定理 1.4.11 (Helly) 任何分布函数列 $\{F_n\}$ 有一个子列 $\{F_{k_n}\}$ 弱收敛.

证明 此定理是点集拓扑中基本的 Tychonov 定理的推论. 在这里给一个简单证明. 取 $\mathbf{R}$ 的一个可列稠子集 $D = \{x_n : n \geqslant 1\}$, 那么 $\{F_n(x_1) : n \geqslant 1\}$ 是有界的, 存在收敛子列 $\{F_{k_n^{(1)}}(x_1)\}$, 假设已取得子列 $\{k_n^{(i)}\}$, 那么因为 $\{F_{k_n^{(i)}}(x_{i+1})\}$ 有界, 我们可以取得 $\{k_n^{(i)}\}$ 的子列 $\{k_n^{(i+1)}\}$, 使得 $\{F_{k_n^{(i+1)}}(x_{i+1})\}$ 收敛. 这样用对角线法, 令 $k_n := k_n^{(n)}$, 那么 $\{F_{k_n}\}$ 在 D 的任何点上都收敛. 事实上, 对任何 $i \geqslant 1$, $\{k_n : n \geqslant i\}$ 是 $\{k_n^{(i)} : n \geqslant i\}$ 的子列, 因此 $\{F_{k_n}(x_i)\}$ 收敛. 推出 $\{F_{k_n}\}$ 弱收敛(见习题). □

最后我们给出测度的淡收敛与弱收敛的定义. 设 μ_n, μ 是 Radon 测度, 如果对任何 $f \in C_c(\mathbf{R})$, $\mu_n(f) \to \mu(f)$, 称 μ_n 淡收敛于 μ, 记 $\mu_n \xrightarrow{V} \mu$. 设 μ_n, μ 是分布, 如果对任何有界连续函数 f 有 $\mu_n(f) \longrightarrow \mu(f)$, 称 μ_n 弱收敛于 μ, 记为 $\mu_n \xrightarrow{w} \mu$. 若 μ_n, μ 是分布, 弱收敛一定淡收敛.

习 题

1. 设 $\{\xi_n\}$ 是随机序列, 满足 $\sum_{n\geqslant 1} \mathbb{E}|\xi_n| < \infty$. 证明: $\sum_{n\geqslant 1} \xi_n$ 几乎处处收敛于一个可积函数 ξ, 且 $\mathbb{E}\xi = \sum_{n\geqslant 1} \mathbb{E}\xi_n$.
2. 随机变量 ξ 可积当且仅当级数 $\sum_{n\geqslant 1} \mathbb{P}(|\xi| \geqslant n) < \infty$.

3. 设 $\{A_n\}$ 是事件列, m 是固定自然数. 用 G 表示至少属于 m 个 A_n 中的元素全体. 证明: G 可测且 $m\mathbb{P}(G) \leqslant \sum\limits_n \mathbb{P}(A_n)$.
4. 设 ξ_n, ξ 是非负可积随机变量, $\xi_n \xrightarrow{p} \xi$, $\mathbb{E}\xi_n \longrightarrow \mathbb{E}\xi$. 证明: $\xi_n \xrightarrow{L^1} \xi$.
5. 分布函数列 F_n 弱收敛当且仅当存在 $\mathbf{R}$ 的一个稠子集 D 使得对任何 $x \in D$, $F_n(x)$ 收敛.
6. 设分布函数 F_n 分别对应于分布 μ_n, 且 F_n 弱收敛, 证明: 对任何 $f \in C_0(\mathbf{R})$, $\mu_n(f)$ 收敛.
7. 设分布函数列 $\{F_n\}$ 弱收敛于一个连续的分布函数 F, 证明: 收敛在 $\mathbf{R}$ 上是一致的.
8. 设 $\{\xi_n\}$ 是随机变量列, 存在 $p > 1$ 使得 $\sup_n \mathbb{E}|\xi_n|^p < \infty$, 证明: $\{\xi_n\}$ 是一致可积的.
9. 证明: (Dunford) 一个一致可积的随机变量族是 L^1- 弱紧的.
10. 设 λ 是 Lebesgue 测度在 $[0,1]$ 上的限制. 对 $[0,1]$ 上的分划列 $\{(0 = x_0^{(n)} < \cdots < x_{j_n}^{(n)} = 1) : n \geqslant 1\}$, 任取 $a_i \in (x_{i-1}^{(n)}, x_i^{(n)}]$, $i = 1, \cdots, j_n$, 定义

$$\mu_n = \sum_{i=1}^{j_n} \left(x_i^{(n)} - x_{i-1}^{(n)} \right) \epsilon_{a_i}.$$

当分划趋于零时, 证明: $\mu_n \xrightarrow{w} \lambda$. 试由此推出 $[0,1]$ 上几乎处处连续的有界 Borel 可测函数是 Rieman 可积的.

11. 设 μ_n, μ 是分布, F_n, F 是对应的分布函数. 证明: $\mu_n \xrightarrow{w} \mu$ 当且仅当 $F_n \xrightarrow{w} F$.
12. 设 $\{\xi_n\}$ 是一致可积的独立随机序列. 证明:

$$\frac{1}{n} \sum_{k=1}^{n} (\xi_k - \mathbb{E}\xi_k) \xrightarrow{L^1} 0.$$

13. 设 $f \in C[0,1]$, 对任何 n, 定义

$$B_n f(x) := \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}.$$

用大数定律证明 $B_n f(x)$ 在 $[0,1]$ 上一致收敛于 f.

14. 设 $X_n \sim \mathrm{N}\left(0, \frac{1}{n}\right)$, 证明: 对任何 $A \in \mathscr{B}$, 有

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X_n \in A) = -\operatorname{ess\,inf}\left\{\frac{|y|^2}{2} : y \in A\right\},$$

其中 ess inf 表示本质下确界.

§1.5 特 征 函 数

设 f 是 $\mathbf{R}^d$ 上有界复值函数, 它关于 d- 维分布 μ 的积分自然地定义为

$$\mu(f) := \mu(\mathrm{Re}f) + \mathrm{i}\mu(\mathrm{Im}f),$$

其中 $\mathrm{Re}f$, $\mathrm{Im}f$ 分别是 f 的实部与虚部. 容易验证许多定理对于复积分依然成立, 特别如控制收敛定理, 及可积条件下的 Fubini 定理.

定义 1.5.1 设 μ 是 d- 维分布. μ 的特征函数定义为

$$\hat{\mu}(x) := \int_{\mathbf{R}^d} \mathrm{e}^{\mathrm{i}x\cdot y}\mu(\mathrm{d}y),$$

其中 $x \cdot y$ 是 $\mathbf{R}^d$ 空间的内积.

分布函数 F (或密度函数 f)的特征函数当然地定义为对应分布的特征函数, 记为 $\widehat{F}$ (或 $\hat{f}$). 如果 μ 是 d- 维随机向量 ξ 的分布, 则 $\hat{\mu}(x) = \mathbb{E}\mathrm{e}^{\mathrm{i}x\cdot\xi}$, 因此 $\hat{\mu}$ 也称为是 ξ 的或对应的分布函数的特征函数. $\mu \mapsto \hat{\mu}$ 实际上是概率测度的 Fourier 变换, 有下列性质:

(1) $\hat{\mu}$ 是 $\mathbf{R}^d$ 上复值函数, $|\hat{\mu}| \leqslant 1$, $\hat{\mu}(0) = 1$;

(2) $\hat{\mu}$ 在 $\mathbf{R}^d$ 上一致连续;

(3) $\overline{\hat{\mu}}(x) = \hat{\mu}(-x)$, $x \in \mathbf{R}^d$;

(4) 设 μ, ν 是两个分布, 则 $\widehat{\mu * \nu} = \hat{\mu} \cdot \hat{\nu}$;

(5) 特征函数是非负定的. 一个 $\mathbf{R}^d$ 上的复值函数 ϕ 称为是非负定的, 如果对任何有限点集 $\{x_1, \cdots, x_m\} \subset \mathbf{R}^d$, 矩阵 $(\phi(x_j - x_k))_{1\leqslant j,\ k\leqslant m}$ 是非负定 Hermite 阵.

(6) 设 μ 是随机变量 ξ 的分布, 如果 $\mathbb{E}|\xi|^n < \infty$, 那么 $\hat{\mu}$ 在 0 点 n 次可导, 且这时有 $\mathrm{i}^n\mathbb{E}\xi^n = \hat{\mu}^{(n)}(0)$.

性质 (1), (3) 直接由定义推出. 为验证 (2), 任取 $t, h \in \mathbf{R}^d$, $|\hat{\mu}(t+h)-\hat{\mu}(t)| \leqslant \int |\mathrm{e}^{\mathrm{i}h\cdot y}-1|\mu(\mathrm{d}y)$, 右侧与 t 无关, 有界收敛定理推出一致连续性. (4) 是 Fubini 定理的直接应用. 为验证 (5), 任取复数 $c_1, \cdots, c_m$,

$$\begin{aligned}\sum_{j,\ k} c_j \hat{\mu}(x_j - x_k)\overline{c_k} &= \int \sum_{j,\ k} c_j \mathrm{e}^{\mathrm{i}x_j\cdot y} \cdot \overline{c_k \mathrm{e}^{\mathrm{i}x_k\cdot y}} \mu(\mathrm{d}y) \\ &= \int \Big|\sum_j c_j \mathrm{e}^{\mathrm{i}x_j\cdot y}\Big|^2 \mu(\mathrm{d}y) \geqslant 0.\end{aligned}$$

而对性质 (6), 由可积性条件, 我们可对特征函数求导得

$$\hat{\mu}^{(n)}(x) = \int_{\mathbf{R}} (\mathrm{i}y)^n \mathrm{e}^{\mathrm{i}xy} \mu(\mathrm{d}y).$$

注意当 n 是偶数时, $\hat{\mu}$ 在 0 点 n 次可导蕴含着 $\mathbb{E}|\xi|^n < \infty$. 而一般情况下不一定成立.

例 1.5.1 对 $a \in \mathbf{R}^d$, 单点分布的特征函数 $\hat{\epsilon}_a(x) = \mathrm{e}^{\mathrm{i}a\cdot x}$.

例 1.5.2 设 μ 是参数为 λ 的 Poisson 分布, 即 $\mu(\{n\}) = \mathrm{e}^{-\lambda}\dfrac{\lambda^n}{n!}$, 则

$$\hat{\mu}(x) = \sum_{n\geqslant 0} \mathrm{e}^{-\lambda} \frac{\lambda^n}{n!} \mathrm{e}^{\mathrm{i}nx} = \mathrm{e}^{-\lambda + \lambda \mathrm{e}^{\mathrm{i}x}}.$$

例 1.5.3 设 ξ 服从 $[a,b]$ 上均匀分布, 即有密度 $f = \dfrac{1}{b-a} 1_{[a,\ b]}$, 则

$$\hat{\mu}_\xi(x) = \frac{1}{b-a}\int_a^b \mathrm{e}^{\mathrm{i}xy}\mathrm{d}y = \frac{\mathrm{e}^{\mathrm{i}bx} - \mathrm{e}^{\mathrm{i}ax}}{\mathrm{i}x(b-a)}.$$

例 1.5.4 设 ξ 服从 $\mathbf{R}$ 上标准正态分布, 即密度为 $f(y) = \dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}}$, 则

$$\begin{aligned}\hat{\mu}_\xi(x) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} \mathrm{e}^{-\frac{y^2}{2} + \mathrm{i}xy} \mathrm{d}y \\ &= \frac{1}{\sqrt{2\pi}} \int \mathrm{e}^{-\frac{y^2 - 2\mathrm{i}xy + (\mathrm{i}x)^2 - (\mathrm{i}x)^2}{2}} \mathrm{d}y \\ &= \frac{1}{\sqrt{2\pi}} \int \mathrm{e}^{-\frac{(y-\mathrm{i}x)^2}{2} - \frac{x^2}{2}} \mathrm{d}y\end{aligned}$$

$$= e^{-\frac{x^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} \int e^{-\frac{(y-ix)^2}{2}} dy$$
$$= e^{-\frac{x^2}{2}}.$$

最后一个积分的计算要用 Cauchy 积分原理: $\int_{-N}^{N} e^{-\frac{(y-ix)^2}{2}} dy$ 可视为解析函数 $z \mapsto e^{-z^2}$ 从 $-N - ix$ 到 $N - ix$ 的(与路径无关)积分, 因此

$$\int_{-N}^{N} e^{-\frac{(y-ix)^2}{2}} dy = \int_{-N-ix}^{N-ix} e^{-\frac{z^2}{2}} dz$$
$$= \int_{-N-ix}^{-N} + \int_{-N}^{N} + \int_{N}^{N-ix} e^{-\frac{z^2}{2}} dz.$$

令 $N \to \infty$, 容易验证最后算式的第一个与第三个积分趋近于零, 而第二个积分是 $\sqrt{2\pi}$. 由此推出密度为 $\frac{1}{(2\pi)^{\frac{d}{2}}} e^{-\frac{|y|^2}{2}}$ 的 d- 维标准正态分布的特征函数是 $x \mapsto e^{-\frac{|x|^2}{2}}$.

由此例容易推出如果 ξ 是服从例 1.3.11 的正态分布的 d 维随机变量, 则其特征函数为

$$x \mapsto \exp\left(i(a,\ x) - \frac{1}{2} x B x^{\mathrm{T}} \right).$$

它当 B 只是非负定时仍然是特征函数, 参见习题. 我们把一个特征函数如上(B 是非负定的)的分布函数也称为正态分布, 即包含了退化的情况. 自然地, 一个随机向量是正态分布的如果其分布函数是正态分布的. 由此可证明一个随机向量是正态分布的当且仅当其分量的任何线性组合是正态分布. ∎

下面我们证明特征函数唯一地决定分布. 设 μ, ν 是 $\mathbf{R}^d$ 上的两个分布, 则对任何 $x, z \in \mathbf{R}^d$, $e^{-ixz}\hat{\mu}(z) = \int e^{iz(y-x)}\mu(dy)$, 那么对变量 z 关于 ν 积分, 由 Fubini 定理得,

$$\int e^{-ixz}\hat{\mu}(z)\nu(dz) = \int \hat{\nu}(y-x)\mu(dy).$$

取 ν 为分布

$$b_t(dx) = \frac{1}{(2\pi t)^{\frac{d}{2}}} e^{-\frac{|x|^2}{2t}} dx,$$

其中 $t>0$, 则由例 1.5.4, $\hat{b}_t(y)=\mathrm{e}^{-t\frac{|y|^2}{2}}$, 因此

$$\int \mathrm{e}^{-\mathrm{i}xz}\hat{\mu}(z)\frac{1}{(2\pi t)^{\frac{d}{2}}}\mathrm{e}^{-\frac{|z|^2}{2t}}\mathrm{d}z=\int \mathrm{e}^{-t\frac{|y-x|^2}{2}}\mu(\mathrm{d}y),\ x\in\mathbf{R}^d.$$

将 t 用 $\dfrac{1}{t}$ 代替,

$$\int \mathrm{e}^{-\mathrm{i}xz}\hat{\mu}(z)\frac{1}{(2\pi)^d}\mathrm{e}^{-\frac{t|z|^2}{2}}\mathrm{d}z=\frac{1}{(2\pi t)^{\frac{d}{2}}}\int \mathrm{e}^{-\frac{|y-x|^2}{2t}}\mu(\mathrm{d}y)=b_t*\mu,\ x\in\mathbf{R}^d.$$

当 $t\downarrow 0$ 时, $b_t*\mu\xrightarrow{w}\mu$ (参考例 1.4.3), 故我们有

$$\mu(f)=\lim_{t\downarrow 0}\frac{1}{(2\pi)^d}\int f(x)\mathrm{d}x\int \mathrm{e}^{-\mathrm{i}xz}\hat{\mu}(z)\mathrm{e}^{-\frac{t|z|^2}{2}}\mathrm{d}z,\ f\in C_0(\mathbf{R}^d),$$

由此推出下面的唯一性定理.

定理 1.5.1　不同的分布有不同的特征函数.

上面的公式里积分与极限一般是不能交换的, 但如果 $\hat{\mu}$ 是 Lebesgue 可积的, 则由控制收敛定理得

$$\mu(f)=\frac{1}{(2\pi)^d}\int f(x)\mathrm{d}x\int \mathrm{e}^{-\mathrm{i}xz}\hat{\mu}(z)\mathrm{d}z,$$

即 μ 是关于 Lebesgue 测度绝对连续的, 密度为

$$x\mapsto\frac{1}{(2\pi)^d}\int_{\mathbf{R}^d}\mathrm{e}^{-\mathrm{i}xz}\hat{\mu}(z)\mathrm{d}z.$$

设 $(\xi_1,\cdots,\xi_n)$ 是 n- 维随机向量, 则它们相互独立等价于对应的 n- 维分布 $\mu_{(\xi_1,\ \cdots,\ \xi_n)}$ 等于 n 个分布的乘积, 即

$$\mu_{(\xi_1,\ \cdots,\ \xi_n)}=\mu_{\xi_1}\times\cdots\times\mu_{\xi_n}.$$

由唯一性定理, 我们有下述定理.

定理 1.5.2　随机变量 $\xi_1,\cdots,\xi_n$ 相互独立当且仅当对任何 $(x_1,\cdots,x_n)\in\mathbf{R}^n$,

$$\mathbb{E}\mathrm{e}^{\mathrm{i}\sum_{k=1}^n x_k\xi_k}=\prod_{j=1}^n\mathbb{E}\mathrm{e}^{\mathrm{i}x_j\xi_j}.$$

最后我们将证明 $\mu \mapsto \hat{\mu}$ 的逆映射是连续的. 设 $\{\mu_n\}$, μ 是分布, 则由定理 1.4.10, $\mu_n \xrightarrow{w} \mu$ 蕴含 $\hat{\mu}_n$ 点点收敛于 $\hat{\mu}$ 且收敛在任何有界区间上是一致的.

定理 1.5.3 设 $\{F_n\}$ 是分布函数列. 若 $\widehat{F}_n$ 点点收敛于一个在零点连续的函数 ϕ, 则 ϕ 是一个分布函数 F 的特征函数且 $F_n \xrightarrow{w} F$.

证明 由 Helly 定理, F_n 有一个子列(不妨仍用 F_n 表示) 弱收敛于某个递增右连续函数 F. 如果 F 是分布函数, 利用证明定理 1.5.1 时所用的恒等式, 对 $x \in \mathbf{R}^d$,

$$\frac{1}{(2\pi)^d}\int \mathrm{e}^{-\mathrm{i}xz}\widehat{F}_n(z)\mathrm{e}^{-\frac{t|z|^2}{2}}\mathrm{d}z = \frac{1}{(2\pi t)^{\frac{d}{2}}}\int \mathrm{e}^{-\frac{|y-x|^2}{2t}}\mathrm{d}F_n(y),$$

因此由 Lebesgue 控制收敛定理和弱收敛性,

$$\frac{1}{(2\pi t)^{\frac{d}{2}}}\int \mathrm{e}^{-\frac{|y-x|^2}{2t}}\mathrm{d}F(y) = \frac{1}{(2\pi)^d}\int \mathrm{e}^{-\mathrm{i}xz}\lim_n \widehat{F}_n(z)\mathrm{e}^{-\frac{t|z|^2}{2}}\mathrm{d}z.$$

因此 F 由 $\widehat{F}_n$ 的极限唯一决定, 推出 F_n 弱收敛于 F.

下面我们只需验证当 ϕ 在 0 点连续时, F 是分布函数就够了. 任取 $t>0$, 由 Fubini 定理,

$$\begin{aligned}
\frac{1}{t}\int_{-t}^{t}(1-\widehat{F}_n(x))\mathrm{d}x &= \frac{1}{t}\int_{-t}^{t}\mathrm{d}x\int_{\mathbf{R}}(1-\mathrm{e}^{\mathrm{i}xy})\mathrm{d}F_n(y)\\
&= \frac{1}{t}\int_{\mathbf{R}}F_n(\mathrm{d}y)\int_{-t}^{t}(1-\mathrm{e}^{\mathrm{i}xy})\mathrm{d}x\\
&= 2\int_{\mathbf{R}}\left(1-\frac{\sin ty}{ty}\right)\mathrm{d}F_n(y)\\
&\geqslant 2\int_{|y|\geqslant\frac{2}{t}}\left(1-\left|\frac{\sin ty}{ty}\right|\right)\mathrm{d}F_n(y)\\
&\geqslant \int_{|y|\geqslant\frac{2}{t}}\mathrm{d}F_n(y).
\end{aligned}$$

由控制收敛定理,

$$\lim_n \frac{1}{t}\int_{-t}^{t}(1-\hat{F}_n(x))\mathrm{d}x = \frac{1}{t}\int_{-t}^{t}(1-\phi(x))\mathrm{d}x.$$

又 ϕ 连续且 $\phi(0)=1$, 故对任何 $\epsilon>0$, 存在 $t>0$, 使得 $\dfrac{1}{t}\int_{-t}^{t}(1-\phi(x))\mathrm{d}x<\epsilon$, 因此存在 N 充分大, 使得当 $n\geqslant N$ 时, $\dfrac{1}{t}\int_{-t}^{t}(1-\widehat{F}_n(x))\mathrm{d}x<\epsilon$, 即得 $\int_{|y|\geqslant\frac{2}{t}}\mathrm{d}F_n(y)<\epsilon$. 等价地 $\int_{|y|<\frac{2}{t}}\mathrm{d}F_n(y)>1-\epsilon$. 弱收敛性推出 $\int_{|y|<\frac{2}{t}}\mathrm{d}F(y)>1-\epsilon$, 因此 F 是分布函数. □

作为应用, 我们来证明著名的中心极限定理. 二项分布近似地是标准正态分布的事实通常称为 DeMoivre-Laplace 中心极限定理, 下面我们应用的方法可以证明稍一般的中心极限定理.

定理 1.5.4 (Lindeberg-Lévy) 设 $\{\xi_n\}$ 是平方可积的独立同分布的随机变量列, $\mathbb{E}\xi_n=a$, $D\xi_n=\sigma^2$, 则 $\dfrac{\sum_{i=1}^{n}(\xi_i-a)}{\sigma\sqrt{n}}$ 的分布弱收敛于标准正态分布.

证明 不妨设 $a=0$, $\sigma=1$, f 是 ξ_n 共同的特征函数, 则 $\dfrac{\sum_{i=1}^{n}\xi_i}{\sqrt{n}}$ 的特征函数为 $t\mapsto f\left(\dfrac{t}{\sqrt{n}}\right)^n$. 由条件知当 $t\to 0$ 时, $f(t)-\left(1-\dfrac{t^2}{2}\right)=o(t^2)$ 即是 t^2 的高价无穷小, 因此对任何 t, 因为 $|f(t)|\leqslant 1$ 且当 n 充分大时, $0\leqslant 1-\dfrac{t^2}{2n}\leqslant 1$, 故

$$\left|f\left(\frac{t}{\sqrt{n}}\right)^n-\left(1-\frac{t^2}{2n}\right)^n\right|\leqslant n\left|f\left(\frac{t}{\sqrt{n}}\right)-\left(1-\frac{t^2}{2n}\right)\right|=n\cdot o\left(\frac{t^2}{n}\right)\longrightarrow 0,$$

推出

$$\lim_n f\left(\frac{t}{\sqrt{n}}\right)^n=\lim_n\left(1-\frac{t^2}{2n}\right)^n=\mathrm{e}^{-\frac{t^2}{2}}.$$

定理结论由定理 1.5.3 推出. □

实际上, 定义 1.5.1 下所列的特征函数性质 (1), (2), (5) 也是特征函数的充分条件. 下面的定理是重要的, 但证明比较难, 而且在此讲义中也基本上用不到, 所以我们省略证明.

定理 1.5.5 (Bochner-Khinchin) 设 ϕ 是 $\mathbf{R}$ 上复值连续函数且 $\phi(0)=1$, 那么 ϕ 是一个概率测度的特征函数当且仅当它是非负定的.

Laplace 变换类似于特征函数, 是研究 $\mathbf{R}_+$ 上测度的重要工具. 设 ξ 是一个非负随机变量, 自然其分布 μ 集中在非负实数集上. 定义

$$L_\xi(t) = L\mu(t) := \mathbb{E}\mathrm{e}^{-t\xi} = \int \mathrm{e}^{-tx}\mu(\mathrm{d}x),\ t \geqslant 0.$$

它实际上是 μ 的 Laplace 变换. 在上下文明确的情况下, 也称它是 ξ 或 μ 的特征函数. 容易验证:

(1) $0 \leqslant L_\xi \leqslant 1$ 且 $L\xi(0) = 1$;

(2) L_ξ 在 $[0,\infty)$ 上连续;

(3) 设 ξ, η 是两个独立随机变量, 则 $L_{\xi+\eta} = L_\xi \cdot L_\eta$;

(4) Laplace 变换作为 t 的函数是完全单调的, 这里 $[0,+\infty)$ 上的函数 f 称为完全单调, 是指 f 是无穷次可微的且 $(-1)^n f^{(n)} \geqslant 0$.

下面是 Laplace 变换的唯一性.

定理 1.5.6 设 μ, ν 是两个 $[0,+\infty)$ 上的分布, 如果 $L\mu = L\nu$, 则 $\mu = \nu$.

证明 把 $[0,\infty)$ Alexander 紧化为 $[0,\infty]$, 那么

$$C[0,\infty] = \{u \in C[0,\infty) : \lim_{x\to\infty} u(x) \text{ 存在}\}.$$

对任何 $t \geqslant 0$, 函数 $x \mapsto \mathrm{e}^{-tx}$ 是属于 $C[0,\infty]$ 的. 用 $\mathscr{A}$ 表示这样的函数的有限线性组合全体, 则 $\mathscr{A}$ 是 $C[0,\infty]$ 的一个区分点的子代数, 由 Stone-Weierstrass 定理得 $\mathscr{A}$ 在 $C[0,\infty]$ 中稠, 而条件说明 μ 与 ν 在 $\mathscr{A}$ 上是一样的, 由稠密性它们在 $C[0,\infty]$ 也一样, 因此 $\mu = \nu$. □

例 1.5.5 设 ξ 服从参数为 λ 的指数分布, 即其密度为 $f(x) = \mathrm{e}^{-\lambda x}1_{\{x\geqslant 0\}}$, 则

$$L_\xi(t) = \int \mathrm{e}^{-(\lambda+t)x}\mathrm{d}x = \frac{1}{\lambda + t}.$$

下面的定理类似于 Bochner-Khinchin 定理.

定理 1.5.7 (Bernstein) 设 ϕ 是 $[0,\infty)$ 上连续函数且 $\phi(0) = 1$, 那么 ϕ 是 $[0,\infty)$ 上的一个概率测度的 Laplace 变换当且仅当它是完全单调的.

最后简单介绍母函数的概念和性质.

定义 1.5.2 设有一个实数列 $a = \{a_0, a_1, \cdots, a_n, \cdots\}$, 如果幂级数

$$G_a(z) := a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n + \cdots$$

在 0 点的一个非空领域上收敛, 则称它为数列 a 的母函数.

由幂级数的知识可以证明, 一个数列的母函数唯一地决定这个数列. 母函数本身只是表达和研究数列的一种方法, 其中的变量 z 没有实质的意义. 设 ξ 是 $\mathbf{Z}_+$-值(甚至可以允许取 $+\infty$)随机变量, 那么它的分布律 $\mathbb{P}(\xi = n)$ 是一个有界数列, 用 G_ξ 表示它的母函数

$$G_\xi(z) := \sum_{n \geqslant 0} z^n \mathbb{P}(\xi = n),$$

那么 $G_\xi(z) = \mathbb{E}(z^\xi)$. 显然 $\lim_{z \uparrow 1} G_\xi(z) = \mathbb{P}(\xi < +\infty)$. 如果 ξ, η 是非负整数值随机变量且 $G_\xi = G_\eta$, 那么 ξ 与 η 有相同的分布律.

设 $\mathbf{Z}_+$- 值随机变量 ξ 的母函数是 G_ξ, 那么

$$\mathbb{E}\xi = G'_\xi(1),\ \mathbb{E}\xi^2 = G''_\xi(1) + G'_\xi(1), \cdots .$$

显然, 结论由下面公式推出

$$G'_\xi(z) = \sum_{n \geqslant 0} n \mathbb{P}(\xi = n) z^{n-1};$$

$$G''_\xi(z) = \sum_{n \geqslant 1} n(n-1) \mathbb{P}(\xi = n) z^{n-2}.$$

定义 1.5.3 数列 $\{u_n\}$ 与 $\{v_n\}$ 的卷积定义为数列

$$\{u_0 v_n + u_1 v_{n-1} + \cdots + u_n v_0 : n \geqslant 0\}.$$

容易验证, 如果 $\mathbf{Z}_+$- 值随机变量 ξ, η 独立, 那么 $\xi + \eta$ 的分布列是 ξ 与 η 的分布律的卷积. 下面的定理也很容易验证, 作为习题.

定理 1.5.8 母函数为 $U(z)$ 的数列 $\{u_n\}$ 与母函数为 $V(z)$ 的数列 $\{v_n\}$ 的卷积的母函数为 $U(z)V(z)$. 因此独立 $\mathbf{Z}_+$- 值的随机变量 ξ, η 的和 $\xi + \eta$ 的母函数

$$G_{\xi+\eta}(z) = G_\xi(z) G_\eta(z).$$

从本质上讲, 特征函数, Laplace 变换与母函数都是类似的, 其中最主要的是它们都唯一地确定分布且将分布的卷积化为乘积.

习　题

1. 一个随机变量 X 称为是格分布的, 如果存在 a 与 $b>0$ 使得 X 支撑在格 $\{a+nb: n=\cdots,-1,0,1,2,\cdots\}$ 上. 设 X 的特征函数为 ϕ.

   (1) X 是格分布的当且仅当存在 $x\neq 0$ 使得 $|\phi(x)|=1$;

   (2) 设存在不可公度的 x,x' (即 $x\neq 0, x'\neq 0, x/x'$ 是无理数) 使得 $|\phi(x)|=|\phi(x')|=1$, 则 X 是常数.

2. 设 ϕ 是一个有密度函数 f 的分布的特征函数,

   (1) 证明: $\lim_{x\to\infty}\phi(x)=0$;

   (2) 证明: (Parseval 等式) 如果 $f\in L^2(\mathbf{R})$, 那么 $\phi\in L^2(\mathbf{R})$ 且

$$\frac{1}{2\pi}\int_{\mathbf{R}}|\phi(t)|^2\mathrm{d}t=\int_{\mathbf{R}}f(x)^2\mathrm{d}x.$$

3. 证明: 例 1.5.4 中函数 $x\mapsto \exp\left(i(a,x)-\dfrac{1}{2}xBx^{\mathrm{T}}\right)$ 当 B 非负定时也是特征函数.

4. d- 维随机向量 $(\xi_1,\cdots,\xi_d)$ 服从正态分布当且仅当 $\xi_1,\cdots,\xi_d$ 的任何线性组合也服从正态分布.

5. (1) 一个服从正态分布的 d- 维随机向量 $(\xi_1,\cdots,\xi_d)$ 独立(即 $\xi_1,\cdots,\xi_d$ 互相独立)当且仅当其协方差矩阵是对角型的.

   (2) 如果 $\boldsymbol{\xi}=(\xi_1,\cdots,\xi_d)$ 服从正态分布, 那么存在正交矩阵 Q 使得随机向量 $\boldsymbol{\xi}Q$ 是独立的.

6. 证明: 一个分布 μ 是对称的 ($\mu(-A)=\mu(A)$, $A\in\mathscr{B}$) 当且仅当特征函数是实值的.

7. 考虑非负实值函数 ϕ 满足 $\phi(-x)=\phi(x)$ 和 $\phi(0)=1$. 证明: (Polya 准则) 如果 ϕ 在 $[0,\infty)$ 上连续且凸, 则 ϕ 是一个特征函数.

8. 设随机变量 X 以概率 1 取无理数, 记 μ_n 是 nX 的小数部分的分布. 证明: $\dfrac{1}{n}\sum\limits_{k=1}^{n}\mu_k$ 弱收敛到 $[0,1]$ 上均匀分布.

9. 设 ξ,η 是两个随机变量, 证明: 如果 ξ 与 η 独立, 则 $\xi+\eta$ 的特征函数等于 ξ 与 η 特征函数的乘积.

10. $\mathbf{R}$ 上密度函数为 $x \mapsto \dfrac{\lambda}{2}\mathrm{e}^{-\lambda|x|}$ 的分布称为是参数为 λ 的 Laplace 分布,计算 Laplace 分布的特征函数.

11. $\mathbf{R}$ 上密度函数为 $x \mapsto \dfrac{1}{\pi}\dfrac{t}{x^2+t^2}$, $t > 0$ 的分布称为是参数为 t 的 Cauchy 分布.

    (1) 计算 Cauchy 分布的特征函数;

    (2) 证明: 两个独立的服从参数分别为 t, s 的 Cauchy 分布的随机变量之和是一个服从参数为 $t+s$ 的 Cauchy 分布.

    (3) 验证如果 ξ 服从参数 1 的 Cauchy 分布, 则 2ξ 的特征函数是 ξ 的特征函数的平方, 以此说明两非独立随机变量和的特征函数可以是两者的乘积.

12. 是否存在独立同分布随机变量 X, Y 使得 $X - Y$ 是 $[-1, 1]$ 上均匀分布?

13. 证明: 当 $q > 2$ 时, $t \mapsto \mathrm{e}^{-|t|^q}$ 不是特征函数.

14. (Bernstein) 设 ξ, η 是独立同分布的平方可积随机变量, 如果 $\xi + \eta$ 与 $\xi - \eta$ 独立, 则 ξ, η 是正态分布的.

15. (Abel) 如果级数 $\sum\limits_n a_n$ 收敛, 那么级数 $\sum\limits_n a_n z^n$ 对所有 $|z| < 1$ 收敛, 且

$$\lim_{z\uparrow 1}\sum_n a_n z^n = \sum_n a_n.$$

16. 设 μ 是 $\mathbf{R}$ 上一个分布. 如果 $\mu(\{a\}) = \mu(\{b\}) = 0$, 证明: 下列公式成立

$$\mu((a,b]) = \frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{\mathrm{e}^{-\mathrm{i}ax} - \mathrm{e}^{-\mathrm{i}bx}}{\mathrm{i}x}\hat{\mu}(x)\mathrm{d}x.$$

因此推出特征函数唯一地确定分布.

§1.6 条件数学期望

条件数学期望是随机分析理论中一个极其重要的概念, 在 Markov 过程和鞅的研究中是不可或缺的. 它与概率论中条件概率的概念有相类似的地方而又有本质的区别. 下面我们先给出其定义.

定义 1.6.1 设 $(\Omega,\mathscr{F},\mathbb{P})$ 是一个概率空间, $\mathscr{A}$ 是 $\mathscr{F}$ 的子 σ- 代数, ξ 是 $(\Omega,\mathscr{F},\mathbb{P})$ 上可积随机变量, ξ 关于 $\mathscr{A}$ 的条件数学期望, 记为 $\mathbb{E}(\xi|\mathscr{A})$, 是指满足以下条件的随机变量 η:

(1) η 是 $\mathscr{A}$ 可测的;

(2) 对任何的 $B\in\mathscr{A}$, $\mathbb{E}(\xi;B)=\mathbb{E}(\eta;B)$. 特别地, 记 $\mathbb{P}(B|\mathscr{A}):=\mathbb{E}(1_B|\mathscr{A})$, 称为 B 关于 $\mathscr{A}$ 的条件概率. 另外, 如果 $\{\xi_i: i\in I\}$ 是一族随机变量. 我们记 $\mathbb{E}(\xi|\xi_i: i\in I):=\mathbb{E}(\xi|\sigma(\xi_i: i\in I))$.

首先我们需要证明条件数学期望的存在性与唯一性. 事实上, 对 $A\in\mathscr{A}$, 令 $\mu(A):=\mathbb{E}(\xi;A)$, 则 μ 是 $(\Omega,\mathscr{A})$ 上的有限符号测度, 再令 $\mathbb{P}_{\mathscr{A}}$ 是 $\mathbb{P}$ 在 $\mathscr{A}$ 上的限制, 那么容易验证 $\mu\ll\mathbb{P}_{\mathscr{A}}$ 且其 Radon-Nikodym 导数满足定义中的条件 (1) 与 (2), 因此是一个条件数学期望. 另外容易知道条件数学期望在几乎处处相等的意义之下是唯一的. 以后如无特别说明, 有关条件数学期望的等式或不等式都是在几乎处处相等的意义之下. 定义中, 加在 ξ 上的可积性条件可以减弱, 有许多书与文章专门讨论此类问题. 但在本书中我们仅讨论可积的情况.

例 1.6.1 设 $\Omega_1,\cdots,\Omega_n$ 是 Ω 的可测分划且 $\mathbb{P}(\Omega_i)>0$, $1\leqslant i\leqslant n$. 令

$$\mathscr{A}:=\sigma(\{\Omega_1,\cdots,\Omega_n\}).$$

取可积随机变量 ξ, 我们来计算 $\mathbb{E}(\xi|\mathscr{A})$. 首先因为 $\mathbb{E}(\xi|\mathscr{A})$ 是 $\mathscr{A}$- 可测的, 故它是一个简单随机变量, 形式为 $\sum\limits_{i=1}^{n}a_i1_{\Omega_i}$. 现在利用定义的条件 (2) 得 $a_i\mathbb{P}(\Omega_i)=\mathbb{E}(\xi;\Omega_i)$, 因此

$$\mathbb{E}(\xi|\mathscr{A})=\sum_{i=1}^{n}\frac{\mathbb{E}(\xi;\Omega_i)}{\mathbb{P}(\Omega_i)}\cdot 1_{\Omega_i}.$$

$\dfrac{\mathbb{E}(\xi;\Omega_i)}{\mathbb{P}(\Omega_i)}$ 称为是 ξ 在 Ω_i 上的平均. ▮

上面的例子是简单情况下条件数学期望的直观解释. 如将 σ- 代数理解为信息, $\mathscr{F}$ 即表示全部的信息. 条件数学期望 $\mathbb{E}(\xi|\mathscr{A})$ 表示在已知信息 $\mathscr{A}$ 下 ξ 的局部平均, 或对 ξ 的某种意义的最好估计. 下面我们给出有关条件数学期望的一些性质.

定理 1.6.1 设 $\xi,\eta,\{\xi_n\}$ 是可积随机变量.

(1) $\mathbb{E}(\xi|\mathscr{F})=\xi$; 如果 ξ 与 $\mathscr{A}$ 独立, 则 $\mathbb{E}(\xi|\mathscr{A})=\mathbb{E}\xi$, 特别地 $\mathbb{E}(\xi|\{\Omega,\varnothing\})=\mathbb{E}\xi$;

(2) 如果 $\xi = a$, 则 $\mathbb{E}(\xi|\mathscr{A}) = a$ a.s.;

(3) 设 a, b 是常数, 则 $\mathbb{E}(a\xi + b\eta|\mathscr{A}) = a\mathbb{E}(\xi|\mathscr{A}) + b\mathbb{E}(\eta|\mathscr{A})$;

(4) 如果 $\xi \leqslant \eta$, 则 $\mathbb{E}(\xi|\mathscr{A}) \leqslant \mathbb{E}(\eta|\mathscr{A})$;

(5) $|\mathbb{E}(\xi|\mathscr{A})| \leqslant \mathbb{E}(|\xi||\mathscr{A})$;

(6) $\mathbb{E}(\mathbb{E}(\xi|\mathscr{A})) = \mathbb{E}\xi$;

(7) 如果 $\lim_n \xi_n = \xi$ a.s. 且 $|\xi_n| \leqslant \eta$, 则

$$\lim_n \mathbb{E}(\xi_n|\mathscr{A}) = \mathbb{E}(\xi|\mathscr{A}).$$

证明　(1), (2), (3) 都是显然的, 为证 (4) 对任何 $A \in \mathscr{A}$,

$$\mathbb{E}([\mathbb{E}(\eta|\mathscr{A}) - \mathbb{E}(\xi|\mathscr{A})]; A) = \mathbb{E}(\mathbb{E}(\eta - \xi|\mathscr{A}); A) = \mathbb{E}((\eta - \xi); A) \geqslant 0.$$

而 $\mathbb{E}(\eta|\mathscr{A}) - \mathbb{E}(\xi|\mathscr{A})$ 是 $\mathscr{A}$ 可测的, 故是非负的. (5) 由 (4) 推出. (6) 是下面定理 1.6.2(2) 的推论. 为证 (7), 令 $Z_n := \sup_{k\geqslant n} |\xi_k - \xi|$, 则 $Z_n \downarrow 0$ 且 $|Z_n| \leqslant 2\eta$, 由控制收敛定理 $\mathbb{E}Z_n \downarrow 0$. 而 $|\mathbb{E}(\xi_n|\mathscr{A}) - \mathbb{E}(\xi|\mathscr{A})| \leqslant \mathbb{E}(Z_n|\mathscr{A})$ 且 $\mathbb{E}(Z_n|\mathscr{A})$ 是单调的, 设其极限是 Z, 则 Z 是非负的, $\mathbb{E}Z = \mathbb{E}(\mathbb{E}(Z|\mathscr{A})) \leqslant \mathbb{E}(\mathbb{E}(Z_n|\mathscr{A})) = \mathbb{E}Z_n$, 因此 $\mathbb{E}Z = 0$, 即 $Z = 0$ a.s.. □

定理 1.6.2　设 ξ, η 是随机变量.

(1) 如果 ξ 是 $\mathscr{A}$ 可测的, 且 η 和 $\xi\eta$ 是可积的, 则 $\mathbb{E}(\xi\eta|\mathscr{A}) = \xi\mathbb{E}(\eta|\mathscr{A})$;

(2) 如果 $\mathscr{A} \subset \mathscr{B}$ 都是 $\mathscr{F}$ 的子 σ- 代数, 且 ξ 是可积的, 则

$$\mathbb{E}(\mathbb{E}(\xi|\mathscr{A})|\mathscr{B}) = \mathbb{E}(\mathbb{E}(\xi|\mathscr{B})|\mathscr{A}) = \mathbb{E}(\xi|\mathscr{A}).$$

证明　(1) 只需对非负的 ξ, η 对就够了. 因 $\xi\mathbb{E}(\eta|\mathscr{A})$ 是 $\mathscr{A}$ 可测的, 故我们只需验证对任意 $A \in \mathscr{A}$, 有

$$\mathbb{E}(\xi\mathbb{E}(\eta|\mathscr{A}); A) = \mathbb{E}(\xi\eta; A).$$

当 ξ 是示性函数时, 即 $\xi = 1_G, G \in \mathscr{A}$, 上式显然成立, 因此上式对 $\mathscr{A}$ 可测的简单函数成立, 从而对非负可测函数成立.

(2) 首先 $\mathbb{E}(\xi|\mathscr{A})$ 是 $\mathscr{B}$ 可测的, 因此必有 $\mathbb{E}(\mathbb{E}(\xi|\mathscr{A})|\mathscr{B}) = \mathbb{E}(\xi|\mathscr{A})$. 另一方面, 对 $A \in \mathscr{A}$, 则 $A \in \mathscr{B}$, 故

$$\int_A \mathbb{E}(\xi|\mathscr{A})\mathrm{d}\mathbb{P} = \int_A \xi \mathrm{d}\mathbb{P} = \int_A \mathbb{E}(\xi|\mathscr{B})\mathrm{d}\mathbb{P}.$$

因此 $\mathbb{E}(\mathbb{E}(\xi|\mathscr{B})|\mathscr{A}) = \mathbb{E}(\xi|\mathscr{A})$. □

这个定理有直观的解释, 实际上是全概率公式的推广, 读者可自己试着去理解. 最后, 我们还有重要的 Jensen 不等式. $\mathbf{R}$ 的区间 (a,b) 上的凸函数 ϕ 是指对任何 $x,y \in (a,b)$ 及 $p,q \geqslant 0$, $p+q=1$, 有

$$\phi(px+qy) \leqslant p\phi(x) + q\phi(y).$$

定理 1.6.3 (Jensen) 设 ξ 是可积随机变量, ϕ 是 $\mathbf{R}$ 上的凸函数且 $\phi(\xi)$ 可积, 则

$$\phi(\mathbb{E}(\xi|\mathscr{A})) \leqslant \mathbb{E}(\phi(\xi)|\mathscr{A}).$$

证明 凸性保证 ϕ 的左右导数存在, 令 A 是其右导数, 则 A 递增且对任何 $x_0 \in \mathbf{R}$, $A(x_0)(x-x_0) + \phi(x_0) \leqslant \phi(x)$, $x \in \mathbf{R}$. 将 x, x_0 分别用 ξ, $\mathbb{E}(\xi|\mathscr{A})$ 代入:

$$A(\mathbb{E}(\xi|\mathscr{A}))[\xi - \mathbb{E}(\xi|\mathscr{A})] + \phi(\mathbb{E}(\xi|\mathscr{A})) \leqslant \phi(\xi).$$

如果 $\mathbb{E}(\xi|\mathscr{A})$ 有界, 则上式左边两项有界, 右边一项可积. 因 A 是 Borel 可测的, 故 $A(\mathbb{E}(\xi|\mathscr{A}))$ 关于 $\mathscr{A}$ 可测. 两边对 $\mathscr{A}$ 取条件数学期望得证 Jensen 公式.

一般地, 令 $G_n := \{\mathbb{E}(|\xi||\mathscr{A}) \leqslant n\}$, 则 $G_n \in \mathscr{A}$ 且 $G_n \uparrow \Omega$. 因此

$$\phi(\mathbb{E}(\xi 1_{G_n}|\mathscr{A})) \leqslant \mathbb{E}(\phi(\xi 1_{G_n})|\mathscr{A}) = \mathbb{E}(1_{G_n}\phi(\xi) + 1_{G_n^c}\phi(0)|\mathscr{A}).$$

由 ϕ 的连续性及控制收敛定理得上述公式. □

例 1.6.2 此例说明条件期望与条件密度的关系. 设 (ξ,η) 是 2 维随机变量, F 是它们的联合分布函数. 设 $x,y \in \mathbf{R}$, 考虑条件概率 $\mathbb{P}(\xi = x|\eta = y)$. 在离散的情况下, 这个概率在 y 属于 η 的值域时是有意义的, 通常记它为 $f_{\xi|\eta}(x|y)$, 那么条件数学期望有表达式 $\mathbb{P}(\xi = x|\eta) = f_{\xi|\eta}(x|\eta)$. 因为, 当 y 属于 η 的值域时,

$$\mathbb{P}(\xi = x, \eta = y) = f_{\xi|\eta}(x|y)\mathbb{P}(\eta = y) = \mathbb{E}(f_{\xi|\eta}(x|\eta); \eta = y).$$

以上的 $x \mapsto f_{\xi|\eta}(x|y)$ 就是 $\eta = y$ 时 ξ 的条件分布律.

一般地, 因为 $\{\eta = y\}$ 可能概率为零, 上面的条件概率无意义. 但是如果 (ξ,η) 是连续型的, 即它们有一个 Borel 可测的密度函数 f, 记 ξ,η 的密度函数分别为

f_ξ, f_η, 那么我们定义

$$\mathbb{P}(\xi \leqslant x|\eta = y) := \lim_{\delta\to 0}\mathbb{P}(\xi \leqslant x|\eta \in (y-\delta, y+\delta)),$$

当然是当右边的极限存在时, 如果极限不存在就定义为 0. 因为

$$\mathbb{P}(\xi \leqslant x|\eta \in (y-\delta, y+\delta)) = \frac{\displaystyle\int_{-\infty}^{x} \mathrm{d}s \int_{(y-\delta,\ y+\delta)} f(s,t)\mathrm{d}t}{\displaystyle\int_{(y-\delta,\ y+\delta)} f_\eta(t)\mathrm{d}t}$$

$$= \frac{\displaystyle\int_{-\infty}^{x} \mathrm{d}s \frac{1}{2\delta}\int_{(y-\delta,\ y+\delta)} f(s,t)\mathrm{d}t}{\displaystyle\frac{1}{2\delta}\int_{(y-\delta,\ y+\delta)} f_\eta(t)\mathrm{d}t},$$

故在几乎处处的意义下,

$$\mathbb{P}(\xi \leqslant x|\eta = y) = \int_{-\infty}^{x} \frac{f(s,y)}{f_\eta(y)}\mathrm{d}s,$$

且当右边的分母为 0 时看作 0. 函数 $\dfrac{f(\cdot,y)}{f_\eta(y)}$ 称为是 $\eta = y$ 时 ξ 的条件概率密度函数, 是 Borel 可测的, 通常记为 $f_{\xi|\eta}(\cdot|y)$. 同样地有表达式

$$\mathbb{P}(\xi \leqslant x|\eta) = \int_{-\infty}^{x} f_{\xi|\eta}(s|\eta)\mathrm{d}s.$$

此式给出了条件期望与条件密度的关系.

更一般地, 因为测度 $B \mapsto \mathbb{P}(\xi \leqslant x, \eta \in B)$ 关于测度 $B \mapsto \mathbb{P}(\eta \in B)$ 绝对连续, 我们将其 Radon-Nikodym 导数用 ϕ 表示, 它是 $\mathbf{R}$ 上的一个几乎处处确定的函数. 显然 $\mathbb{P}(\xi \leqslant x|\eta) = \phi(\eta)$. 当然如果 $\mathbb{P}(\xi \leqslant x|\eta = y)$ 有意义时, 它与 $\phi(y)$ 对于几乎所有的 $y \in \mathbf{R}$ 是相同的. ∎

习　题

1. 设 $\Omega = [-1,1]$, $\mathscr{F}$ 是 Ω 上 Borel 集全体, $\mathbb{P}$ 是 $\dfrac{1}{2}$ 的 Lebesgue 测度, $Y(\omega) = |\omega|$, X 是可积随机变量, 计算: $\mathbb{E}(X|Y)$.

2. 设 (X,Y) 服从参数为 $(a_1,a_2,\sigma_1^2,\sigma_2^2,\rho)$ 的二维正态分布, 求 $\mathbb{E}(X|Y)$.
3. 设 ξ 是随机变量, 证明: ξ 与子 σ- 代数 $\mathscr{A}$ 独立当且仅当对任何有界 Borel 可测函数 g, 有 $\mathbb{E}(g(\xi)|\mathscr{A})=\mathbb{E}g(\xi)$.
4. 设随机变量 X,Y 与 Z 独立, X 可积, 证明: $\mathbb{E}(X|Y,Z)=\mathbb{E}(X|Y)$.
5. 设 ξ,η 是独立同分布的可积随机变量, 证明:

$$\mathbb{E}(\xi|\xi+\eta)=\frac{\xi+\eta}{2}.$$

6. 设 $\xi_1,\xi_2,\cdots$ 是独立同分布的可积随机序列, $S_n:=\xi_1+\cdots+\xi_n$. 证明:

$$\mathbb{E}(\xi_1|S_n,S_{n+1},\cdots)=\frac{S_n}{n}.$$

7. 设 X 服从参数为 (a,σ^2) 的正态分布, Φ 是标准正态分布函数, 求 $\mathbb{E}\Phi(X)$. (提示: 取独立于 X 的服从标准正态分布的随机变量 U, 证明: $\Phi(X)=\mathbb{P}(U\leqslant X|X)$.)
8. 对有界的随机变量 ξ,η 证明: $\mathbb{E}(\xi\mathbb{E}(\eta|\mathscr{A}))=\mathbb{E}(\eta\mathbb{E}(\xi|\mathscr{A}))$.
9. 设 (X,Y) 是 2-维有界随机变量, X 是 $\mathscr{A}$ 可测的, Y 独立于 $\mathscr{A}$. 对 $B\in\mathscr{B}(\mathbf{R}^2)$, 定义 $f_B(x):=\mathbb{P}((x,Y)\in B)$. 证明: $\mathbb{P}((X,Y)\in B|\mathscr{A})=f_B(X)$.
10. 设 $(\Omega,\mathscr{F},\mathbb{P})$ 是概率空间 $(\Omega_1,\mathscr{F}_1,\mathbb{P}_1)$ 与 $(\Omega_2,\mathscr{F}_2,\mathbb{P}_2)$ 的乘积概率空间. 记 $\mathscr{A}:=\{A_1\times\Omega_2:A_1\in\mathscr{F}_1\}$. 对 Ω 上随机变量 f, 定义

$$g(\omega_1):=\int_{\Omega_2}f(\omega_1,\omega_2)\mathbb{P}_2(\mathrm{d}\omega_2).$$

证明: $g=\mathbb{E}(f|\mathscr{A})$ a.s.

11. 设 $\mathscr{A}$ 是 σ- 代数, 非空集 $A\in\mathscr{A}$ 称为是 $\mathscr{A}$ 的原子, 如果 A 没有非平凡子集在 $\mathscr{A}$ 中. 证明 $\mathbb{E}(\xi|\mathscr{A})$ 在 A 上是常数.
12. 设 $\sum$ 是 $\mathscr{F}$ 的子 σ- 代数全体, ξ 是可积随机变量, 证明: $\{\mathbb{E}(\xi|\mathscr{A}):\mathscr{A}\in\Sigma\}$ 是一致可积的.
13. 设 X,Y 独立可积且 $\mathbb{E}X=\mathbb{E}Y=0$, 证明: $\mathbb{E}|X|\leqslant\mathbb{E}|X+Y|$.
14. (1) 如果 $\xi\in L^2(\Omega,\mathscr{F},\mathbb{P})$, 令 $M:=L^2(\Omega,\mathscr{A},\mathbb{P})$, 证明: $\mathbb{E}(\xi|\mathscr{A})$ 是 ξ 在闭子空间 M 上的正交投影;

    (2) 利用 $L^2(\Omega,\mathscr{F},\mathbb{P})$ 在 $L^1(\Omega,\mathscr{F},\mathbb{P})$ 中稠密, 证明条件数学期望的存在性.

第二章　随机过程基础

在这一章中, 我们将首先介绍随机过程的理论定理: Kolmogorov 相容定理, 这实际上是一个特殊的测度扩张定理, 是在无穷维空间上构造测度的方法. 另外我们将介绍一些重要的随机过程, 主要是 Markov 过程, 如 Markov 链, Lévy 过程, Poisson 过程与 Brown 运动. 期望初学者在了解这些具体的随机过程的同时, 对概率论及一般随机过程理论有一个系统而直观的认识.

§2.1　随机过程与无穷乘积空间上的测度

设 T 是一个指标集, 它可以是任意的, 但在本书中, 我们一般取 T 是非负整数集 $\mathbf{Z}^+ = \{0, 1, 2, \cdots\}$ 或非负实数集 $\mathbf{R}^+ = [0, \infty)$ 或它们的子集, 分别称为离散时间集与连续时间集. 下面当我们说时间集时, 是指两者之一. 实际上在许多情况下, 也可以是 $\mathbf{R}$ 的一个子集.

定义 2.1.1　设 $(\Omega, \mathscr{F}, \mathbb{P})$ 是一个概率空间, $(E, \mathscr{E})$ 是一个可测空间, 则一个取值在 E 上的可测映射族 $X = (X_t : t \in \mathsf{T})$ 称为是一个 $(\Omega, \mathscr{F}, \mathbb{P})$ 上以 $(E, \mathscr{E})$ 为状态空间的随机过程. 在上下文明确时, 简称为 E-值随机过程或随机过程. 当 E 是实数空间或复数空间时, 分别称过程是实值过程与复值过程. 特别地, 如果需要, 我们可以把随机变量, 随机向量看成为随机过程.

例 2.1.1　最简单也最早被人们研究的随机过程是随机游动. 设 $\{\xi_n : n \geqslant 1\}$ 是某个概率空间上独立同分布的随机变量序列且都服从 Bernoulli 分布, 即

$$\mathbb{P}(\xi_n = 1) = p,\ \mathbb{P}(\xi_n = -1) = q,$$

其中 $p, q \geqslant 0,\ p + q = 1$. 称这样的随机序列为 Bernoulli (随机)序列. Bernoulli 序列的存在性在直观上是显然的, 但我们将在下一节证明. 这相当于甲乙两人用某种

固定的方法与规则进行一系列独立的赌博. 无疑 Bernoulli 序列 $\{\xi_n : n \geqslant 1\}$ 是一个随机过程, 但更有意思的是下面的过程. 让值 1 表示甲赢, 这时他得到 1 元钱; 值 -1 表示甲输, 这时他付出 1 元钱. 记 S_n 为 n 次赌博后, 甲所拥有的赌资. 任取整数值随机变量 S_0 并令 $S_n = S_0 + \sum\limits_{i=1}^{n} \xi_i,\ n \geqslant 1$. 自然 $\{S_n : n \geqslant 0\}$ 也是一个随机过程, 称为是随机游动. 如果 $S_0 = x \in \mathbf{Z}$, 称 $\{S_n\}$ 是从 x 出发的随机游动. ▮

例 2.1.2 设 $\Omega = [0,1]$, $\mathscr{F} = \mathscr{B}[0,1]$, $\mathbb{P}$ 是 $[0,1]$ 上 Lebesgue 测度. 对 $t \in [0,1]$, 令 $X_t(\omega) := 1_{[0,t]}(\omega)$, 则 $(X_t : t \in [0,1])$ 是一个随机过程. ▮

随机过程就是随机变量的集合, 随机过程的分布是由有限维分布族来描述. 一个具有给定分布的随机过程的存在性远不是一件平凡的事情, 它是由 Kolmogorov 在20世纪30年代证明的, 是随机过程的理论基础. 首先让我们引入随机过程的有限维分布族的概念. 下面描述随机过程时, 通常要涉及较多的符号, 我们期望读者能熟悉下面的符号, 以帮助理解随机过程的概念.

设 $(\Omega, \mathscr{F}, \mathbb{P})$ 是概率空间, $X = (X_t : t \in \mathsf{T})$ 是其上的以 $(E, \mathscr{E})$ 为状态空间的随机过程. 用 $\mathscr{I}_{\mathsf{T}}$ 表示 T 的有序的有限子集全体, 即

$$\mathscr{I}_{\mathsf{T}} := \{(t_1, \cdots, t_n) :\ n \geqslant 1,\ t_1, \cdots, t_n \in \mathsf{T}\}.$$

对 $I = (t_1, \cdots, t_n) \in \mathscr{I}_{\mathsf{T}}$, 记 E^n 为 E^I, X_I 或 $X_{(t_1, \cdots, t_n)}$ 是映射

$$\omega \mapsto (X_{t_1}(\omega), \cdots, X_{t_n}(\omega)),$$

则 X_I 是 Ω 到 E^I 的可测映射. 用 μ_I 表示空间 $(E^I, \mathscr{E}^I)$ 上由 X_I 诱导的测度 $\mathbb{P} \circ X_I^{-1}$, 即 $\mathbb{P}$ 在 X_I 下的像测度. 确切地, 对任何 $A_1, \cdots, A_n \in \mathscr{E}$,

$$\mu_{(t_1, \cdots, t_n)}(A_1 \times \cdots \times A_n) = \mathbb{P}(X_{t_1} \in A_1, \cdots, X_{t_n} \in A_n).$$

再记 $\mathscr{D}_X := \{\mu_I :\ I \in \mathscr{I}_{\mathsf{T}}\}$, 它称为随机过程 X 的有限维分布族.

定理 2.1.1 随机过程 X 的有限维分布族由过程唯一决定且满足下面的性质:

(1) 横向相容: 对于 $I = (t_1, \cdots, t_n) \in \mathscr{I}_{\mathsf{T}}$ 及 $(1, 2, \cdots, n)$ 的一个置换 σ, 记 $\sigma(I) := (t_{\sigma(1)}, \cdots, t_{\sigma(n)})$, 则

$$\mu_{\sigma(I)} = \mu_I \circ \sigma^{-1},$$

上式中 σ 表示由置换诱导的 E^I 上坐标置换:

$$\sigma : (x_1, \cdots, x_n) \mapsto (x_{\sigma(1)}, \cdots, x_{\sigma(n)}).$$

(2) 纵向相容: 设 $I = (t_1, \cdots, t_n) \in \mathscr{I}_{\mathsf{T}}$, $B_1, \cdots, B_n \in \mathscr{E}$, 如果对某个 $1 \leqslant k \leqslant n$ 有 $B_k = E$, 则

$$\begin{aligned}&\mu_I(B_1 \times \cdots \times B_k \times \cdots \times B_n)\\&= \mu_{I_k}(B_1 \times \cdots \times B_{k-1} \times B_{k+1} \times \cdots \times B_n),\end{aligned}$$

其中 $I_k := (t_1, \cdots, t_{k-1}, t_{k+1}, \cdots, t_n)$.

证明　因为 $\{X_t \in E\} = \Omega$, 故 (2) 由定义容易验证. 对于 (1), 也是由定义

$$\begin{aligned}&\mu_{\sigma(I)}(A_1 \times \cdots \times A_n)\\&\quad= \mathbb{P}(X_{t_{\sigma(1)}} \in A_1, \cdots, X_{t_{\sigma(n)}} \in A_n)\\&\quad= \mathbb{P}(X_{t_1} \in A_{\sigma^{-1}(1)}, \cdots, X_{t_n} \in A_{\sigma^{-1}(n)})\\&\quad= \mu_I(A_{\sigma^{-1}(1)} \times \cdots \times A_{\sigma^{-1}(n)})\\&\quad= \mu_I(\sigma^{-1}(A_1 \times \cdots \times A_n)).\end{aligned}$$

完成证明. □

当 T 是全序集时, 横向相容是不那么重要的, 因为对任何 $I = (t_1, \cdots, t_n) \in \mathscr{I}_{\mathsf{T}}$, 我们总可以假设它可以排成 $t_1 < \cdots < t_n$.

后面我们将看到过程的有限维分布族是研究随机过程的一个很好的工具, 有时它比随机过程本身更为重要. 一般说来, 不同的随机过程可能有相同的有限维分布族. 因此我们给出下面的定义.

定义 2.1.2　(1) 分别在概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 和 $(\Omega', \mathscr{F}', \mathbb{P}')$ 上且有相同的状态空间 $(E, \mathscr{E})$ 与相同的指标集 T 的随机过程 X, X' 称为是等价的, 如果它们有相同的有限维分布族, 即对任何 $I \in \mathscr{I}_{\mathsf{T}}$, $X_I(\mathbb{P}) = X'_I(\mathbb{P}')$.

(2) 两个定义在同一个概率空间上具同一个状态空间的随机过程 X, X' 称为是互为修正, 如果对任何 $t \in \mathsf{T}$, $X_t = X'_t$ a.s.. 它们称为是不可区别的, 如果对几乎所有的 $\omega \in \Omega$, $X_t(\omega) = X'_t(\omega)$ 对所有的 t 成立.

显然, 如果 X 和 X' 是不可区别的, 则它们互为修正, 而若 X 和 X' 互为修正, 则它们一定是等价的.

例 2.1.3 设 $\mathbb{P}$ 是 $([0,1],\mathscr{B}([0.1]))$ 上的 Lebesgue 测度, 对 $t,\omega\in[0,1]$, 令 $X_t(\omega)=0$, $X'_t(\omega)=1_{\{t\}}(\omega)$, 则 (X_t) 与 (X'_t) 互为修正, 但非不可区别的.

设 T 是 $\mathbf{R}$ 的一个区间, E 是一个度量空间, $X=(X_t:t\in\mathsf{T})$ 是 $(\Omega,\mathscr{F},\mathbb{P})$ 上以 E 为状态空间的随机过程. 如果当 $s\to t$ 时, X_s 以概率收敛于 X_t, 则称过程 X 在 $t\in\mathsf{T}$ 处随机连续. 如果它在任意点 $t\in\mathsf{T}$ 处随机连续, 则称 X 随机连续. 显而易见上例中的两个过程都是随机连续的. 上面对于随机过程的考虑是将它们作为个别随机变量的一个集合而已, 更有意义的是将它们作为一个整体考虑, 即考虑样本轨道. 对任何 $\omega\in\Omega$, $t\mapsto X_t(\omega)$ 是 T 到 E 的映射, 称为是 ω 的样本轨道. 设 E 是一个拓扑空间, 我们通常要考虑轨道的某种正则性. 称随机过程 X 是右连续(对应地, 左连续, 右连续并存在左极限, 等对应性质)的, 如果其几乎所有样本轨道是右连续(对应地, 左连续, 右连续并存在左极限, 等对应性质)的. 显然连续的过程是随机连续的, 反过来, 在上面的例 2.1.3 中, X 是一个连续过程, X' 不是, 因为它的所有样本轨道都在某个点间断. 因此我们可以看出, 过程的随机连续与连续有本质的不同. 在后面, 我们将分别讨论 Poisson 过程, Brown 运动, 鞅等过程的轨道正则性问题.

随机过程的有限维分布的重要性主要表现在理论方面, 在实际问题中, 两个更容易计算的量是均值函数与协方差函数. 一个复或实值随机过程 $X=(X_t:t\in\mathsf{T})$ 称为可积的, 如果对任何 $t\in\mathsf{T}$, $\mathbb{E}|X_t|<\infty$, 这时 $m(t):=\mathbb{E}X_t$ 称为是过程的均值函数; 称为是平方可积的, 如果对任何 $t\in\mathsf{T}$, $\mathbb{E}|X_t|^2<\infty$; 称为是一致可积的, 如果 $\{X_t:t>0\}$ 是一致可积的随机变量族; 称为是有界的, 如果存在常数 $C>0$, 使得 $|X_t(\omega)|\leqslant C$ 对任何 t,ω 成立. 若 X 是平方可积过程, 我们定义过程的协方差函数

$$K(t,s):=\operatorname{cov}(X_t,X_s)=\mathbb{E}\overline{(X_t-\mathbb{E}X_t)}(X_s-\mathbb{E}X_s),\ t,s\in\mathsf{T}.$$

容易验证协方差函数满足下列性质:

(1) K 是共轭对称的, 即 $K(s,t)=\overline{K(t,s)}$, $s,t\in\mathsf{T}$;

(2) K 是非负定的, 即对任何 $c_1,\cdots,c_n\in\mathbf{C}$, $t_1,\cdots,t_n\in\mathsf{T}$, 有

$$\sum_{j,k}\overline{c_j}K(t_j,t_k)c_k\geqslant 0.$$

例 2.1.4 (Gauss 过程)　概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的实值随机过程 $X = (X_t : t \in \mathsf{T})$ 称为是 Gauss 过程, 如果它的每一个有限维分布都是 Gauss 分布. 进一步, 一个 Gauss 过程称为是中心化的 Gauss 过程, 如果 $\mathbb{E}X_t = 0$, $t \in \mathsf{T}$. 设 $(\Omega, \mathscr{F}, \mathbb{P})$ 是概率空间, ξ 是服从标准正态分布的随机变量, 对 $t \geqslant 0$, 令 $X_t = \xi t$, 则 (X_t) 是一个平方可积的随机过程且 $\mathbb{E}X_t^2 = t^2$, $K(t,s) = ts$. 另外对 $t_1, \cdots, t_n \in \mathsf{T}$, 有限维分布是一个退化的正态分布, 其特征函数是

$$f_{t_1,\cdots,t_n}(x) = \exp\big(-\frac{1}{2}xBx^{\mathrm{T}}\big),$$

其中 $x \in \mathbf{R}^n$, $B = (t_1, \cdots, t_n)^{\mathrm{T}}(t_1, \cdots, t_n)$. ▮

例 2.1.5 (平稳过程)　概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的状态空间为 E 的随机过程 $X = (X_t : t \in \mathsf{T})$ 称为是平稳过程, 如果其有限维分布族是平移不变的. 精确地说 T 是一个时间半群, 即对加法封闭, 且对任何 $t_1, \ldots, t_n, t \in \mathsf{T}$, $A_1, \ldots, A_n \in \mathscr{E}$, 有

$$\mathbb{P}(X_{t_1+t} \in A_1, \ldots, X_{t_n+t} \in A_n) = \mathbb{P}(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n).$$

这也常称为是严格的平稳过程. ▮

例 2.1.6 (独立增量过程)　设 $\mathsf{T} = [0, \infty)$ 或 $\{0, 1, 2, \cdots\}$, $X = (X_t : t \in \mathsf{T})$ 是一个以局部紧 Abel 加群 $(G, +)$ 为状态空间的随机过程, 如果对任何 $0 \leqslant t_1 < \cdots < t_n$, $X_{t_n} - X_{t_{n-1}}, \cdots, X_{t_2} - X_{t_1}, X_{t_1}$ 相互独立, 称 X 是一个独立增量过程. 令 $\mathscr{F}_t := \sigma(\{X_s : s \leqslant t\})$, 包含的是过程的过去直到时刻 t 的信息. 因为 $X_{t_n} - X_{t_{n-1}}, \cdots, X_{t_2} - X_{t_1}, X_{t_1}$ 与 $X_{t_n}, \cdots, X_{t_2}, X_{t_1}$ 可以相互线性表示, 故 X 是独立增量当且仅当对任何 $t > s$, $X_t - X_s$ 与 $\mathscr{F}_s$ 独立. 如果进一步对任何 $t > s$, $X_t - X_s$ 与 $X_{t-s} - X_0$ 同分布, 称 X 是平稳独立增量过程的或空间齐次过程. 我们来计算独立增量过程 X 的有限维分布族. 对 $t > s \geqslant 0$, $A \in \mathscr{B}(G)$, 令 $\mu_{s,t}(A) := \mathbb{P}(X_t - X_s \in A)$, $\mu(A) := \mathbb{P}(X_0 \in A)$, $A \in \mathscr{B}(\mathbf{R}^d)$, 则对 $0 = t_0 < t_1 < \cdots < t_n$, $(X_{t_n} - X_{t_{n-1}}, \cdots, X_{t_1} - X_{t_0}, X_{t_0})$ 的联合分布是

$$\mu_{t_{n-1},t_n} \times \cdots \times \mu_{t_0,t_1} \times \mu.$$

因此对 $A_0, A_1, \cdots, A_n \in \mathscr{B}(\mathbf{R}^d)$,

$$\mathbb{P}(X_{t_0} \in A_0, X_{t_1} \in A_1, \cdots, X_{t_n} \in A_n)$$

$$= \int_{x_0+\cdots+x_i\in A_i,0\leqslant i\leqslant n} \mu(\mathrm{d}x_0)\mu_{t_0,t_1}(\mathrm{d}x_1)\cdots\mu_{t_{n-1},t_n}(\mathrm{d}x_n)$$

$$= \int_{y_i\in A_i,0\leqslant i\leqslant n} \mu(\mathrm{d}y_0)\mu_{t_0,t_1}(\mathrm{d}y_1-y_0)\cdots\mu_{t_{n-1},t_n}(\mathrm{d}y_n-y_{n-1}),$$

其中最后一个等式由变量替换:

$$y_i = x_0 + x_1 + \cdots + x_i,\ i = 0, 1, \cdots, n$$

完成, 其中对任何测度 μ, 符号 $\mu(\mathrm{d}y - x)$ 表示测度 $A \mapsto \mu(A - x)$. 另外容易验证 X 是平稳独立增量过程当且仅当对任何 $s \geqslant 0$, 过程 $t \mapsto X_{t+s} - X_s$ 与 $\mathscr{F}_s$ 独立且与过程 $t \mapsto X_t - X_0$ 等价. ▮

例 2.1.7 (广义平稳过程与正交增量过程) 设 $\mathsf{T} = \mathbf{R}$ (或 $\mathsf{T} = \mathbf{Z}$). 一个复值平方可积过程 $(X_t : t \in \mathsf{T})$ 称为是 L^2-连续的, 如果过程 X 看作 T 到 $L^2(\Omega, \mathscr{F}, \mathbb{P})$ 的映射是连续的. 一个 L^2-连续的随机过程称为广义平稳过程, 在这里我们简称为平稳过程, 如果其均值函数是常数且协方差函数 $K(t,s)$ 只与 $s-t$ 有关, 即存在 T 上函数 K, 使得 $K(t,s) = K(s-t)$, $s,t \in \mathsf{T}$. 那么我们有

(1) $+\infty > K(0) > 0$;

(2) K 是共轭对称的, 即 $K(-t) = \overline{K(t)}$, $t \in \mathsf{T}$;

(3) K 是非负定的, 即对任何 $c_1, \cdots, c_n \in \mathbf{C}$, $t_1, \cdots, t_n \in \mathsf{T}$, 有

$$\sum_{j,k} \overline{c_j} K(t_j - t_k) c_k \geqslant 0.$$

由 Bochner-Khinchin 定理, 存在 $\mathbf{R}$ 上有限测度 μ, 使

$$K(t) = \int_{\mathbf{R}} \mathrm{e}^{\mathrm{i}tx} \mu(\mathrm{d}x).$$

显然过程是实值的当且仅当 μ 是对称的. (如果 $\mathsf{T} = \mathbf{Z}$, 则 K 是一个序列, 这时 μ 将集中在区间 $[-\pi, \pi)$ 上, 并且以下的讨论是平行的.) 若令 $F(x) := \mu((-\infty, x])$, $x \in \mathbf{R}$, 则 F 是单增右连续有界函数, 且 $F(-\infty) = 0, F(+\infty) = K(0)$. μ 和 F 由 K 唯一决定, 分别称为是平稳过程 X 的谱测度与谱函数. μ 有密度时, 密度函数称为是 X 的谱密度.

为了进一步讨论过程的谱表示, 我们首先介绍正交增量过程. 以 T 为指

标集的复值平方可积随机过程 $Z = (Z_x : x \in \mathsf{T})$ 称为是正交增量过程, 如果 $\mathbb{E}Z_x = 0$, $x \in \mathbf{R}$ 且对任何 $x_1 \leqslant x_2 \leqslant x_3 \leqslant x_4$, 有

$$\mathbb{E}(\overline{Z_{x_4} - Z_{x_3}})(Z_{x_2} - Z_{x_1}) = 0.$$

设 Z 是一个正交增量过程, 则存在 T 上单调增加的函数 λ_Z 使得对任何 $y > x$,

$$\mathbb{E}|Z_y - Z_x|^2 = \lambda_Z(y) - \lambda_Z(x).$$

事实上, 由正交性, 对任何 $x_1 \leqslant x_2 \leqslant x_3$, $\mathbb{E}|Z_{x_3} - Z_{x_1}|^2 = \mathbb{E}|Z_{x_3} - Z_{x_2}|^2 + \mathbb{E}|Z_{x_2} - Z_{x_1}|^2$. 因此, 函数 $\lambda_Z(x) := \mathrm{sgn}(x)\mathbb{E}|Z_x - Z_0|^2$, $x \in \mathsf{T}$ 满足要求.

满足上述条件的 λ_Z 在相差一个常数不计的条件下是唯一的. 它称为是正交增量过程 Z 的指标函数. 由它决定的测度 μ_Z 称为是指标测度. 现在我们回到广义平稳过程 X, 不妨设 $m = \mathbb{E}X_t = 0$, (否则, 考虑 $X'_t = X_t - m$.) 令

$$\mathscr{G} := \left\{ f(x) = \sum_{j=1}^{n} c_j \mathrm{e}^{\mathrm{i}xt_j} : \ n \geqslant 1, \ c_1, \cdots, c_n \in \mathbf{C}, t_1, \cdots, t_n \in \mathsf{T} \right\},$$

那么 $\mathscr{G}$ 在 $L^2(\mu)$ 中是稠线性子空间(注意这个事实并不是平凡的). 对 $f : x \mapsto \sum\limits_{j=1}^{n} c_j \mathrm{e}^{\mathrm{i}xt_j}$, $g : x \mapsto \sum\limits_{j=1}^{m} b_j \mathrm{e}^{\mathrm{i}xs_j}$, 令

$$F := \sum_{j=1}^{n} c_j X_{t_j}, \ G := \sum_{j=1}^{m} b_j X_{s_j},$$

则 $f, g \in \mathscr{G}$, 而 $F, G \in L^2(\Omega, \mathscr{F}, \mathbb{P})$. 计算 F, G 的 L^2-内积,

$$\begin{aligned}
(F, G)_{L^2(\mathbb{P})} &= \mathbb{E}\overline{\sum_{j=1}^{n} c_j X_{t_j}} \sum_{j=1}^{m} b_j X_{s_j} \\
&= \sum_{j=1}^{n} \sum_{k=1}^{m} \overline{c_j} b_k K(s_k - t_j) \\
&= \int_{\mathbf{R}} \sum_{j=1}^{n} \sum_{k=1}^{m} \overline{c_j} b_k \mathrm{e}^{\mathrm{i}x(s_k - t_j)} \mu(\mathrm{d}x) \\
&= \int_{\mathbf{R}} \overline{f(x)} g(x) \mu(\mathrm{d}x) = (f, g)_{L^2(\mu)}.
\end{aligned}$$

由上述结论, 映射 $I: f \mapsto F$ 是良定义的, 且是 $\mathscr{G} \subset L^2(\mu)$ 到 $L^2(\mathbb{P})$ 的线性的保持内积不变的一个映射, 它可以连续地延拓到 $L^2(\mu)$ 上, 并且仍然是线性的和保持内积不变. 现在对 $x \in \mathsf{T}$, 令 $Z_x := I(1_{(-\infty,x]})$, 则 $Z = (Z_x : x \in \mathsf{T})$ 是一个平方可积的复值过程, 注意 Z_x 是几乎处处相等的意义下唯一决定的. 由 I 的性质, Z 是一个期望为零的正交增量过程, 且其指标函数恰是 F.

反过来, 设 $Z = (Z_x :\ x \in \mathsf{T})$ 是复值正交增量过程且其指标测度 μ_Z 有限. 令

$$\mathscr{G}_0 := \left\{ \sum_{j=1}^{n} c_j 1_{(x_{j-1},x_j]} :\ n \geqslant 1, x_0 < \cdots < x_n, c_1, \cdots, c_n \in \mathbf{C} \right\}.$$

对 $f \in \mathscr{G}_0$ 有表示 $f = \sum\limits_{j=1}^{n} c_j 1_{(x_{j-1},x_j]}$, 定义

$$I(f) := \sum_{j=1}^{n} c_j (Z_{x_j} - Z_{x_{j-1}}).$$

不难验证, 定义无歧义. 再取 $g \in \mathscr{G}_0$ 有表示 $g = \sum\limits_{j=1}^{m} b_j 1_{(x_{j-1},x_j]}$. 这里我们有理由设 f, g 有同样的分点, 因为否则, 我们可以加细分点. 现在类似地计算 $I(f)$, $I(g)$ 在 $L^2(\Omega, \mathscr{F}, \mathbb{P})$ 中的内积

$$\begin{aligned}(I(f), I(g))_{L^2(\mathbb{P})} &= \sum_{j,k=1}^{n} \overline{c_j} b_k \mathbb{E}[\overline{(Z_{x_j} - Z_{x_{j-1}})}(Z_{x_k} - Z_{x_{k-1}})] \\ &= \sum_{j=1}^{n} \overline{c_j} b_j [\lambda_Z(x_j) - \lambda_Z(x_{j-1})] = (f, g)_{L^2(\mu_Z)}.\end{aligned}$$

因 $\mathscr{G}_0$ 在 $L^2(\mu_Z)$, 同理可以将 I 延拓为 $L^2(\mu)$ 到 $L^2(\mathbb{P})$ 的线性的和保持内积不变映射. 令 $X_t := I(\mathrm{e}^{\mathrm{i}xt})$, 则 $\mathbb{E}X_t = 0$, 且

$$\mathbb{E}(\overline{X_t} X_s) = (I(\mathrm{e}^{\mathrm{i}xt}), I(\mathrm{e}^{\mathrm{i}xs}))_{L^2(P)} = \int_{\mathsf{T}} \mathrm{e}^{-\mathrm{i}x(t-s)} \mu(\mathrm{d}x).$$

因此 $X = (X_t :\ t \in \mathbf{R})$ 是平稳过程且其谱测度是 μ.

我们经常也写 $I(f)$ 为积分形式 $\int f \mathrm{d}Z$, 称为是 f 关于正交增量过程 Z 的积分. 这个积分并不是通常 Stieltjes 意义下的积分, 事实上轨道 $x \mapsto Z_x(\omega)$ 一般地不是有

界变差的, 称它为积分是因为映射 I 与积分有着完全类似的性质. 总结上面讨论, 我们立即得到下列定理.

定理 2.1.2　谱测度为 μ, 且零均值的广义平稳过程 X 与指标测度为 μ 的正交增量过程一一对应. 更多地, 存在一个等距嵌入映射 $I : L^2(\mathbf{R}, \mathscr{B}, \mu) \longrightarrow L^2(\Omega, \mathscr{F}, \mathbb{P})$, 使得 $Z_x = I(1_{(-\infty,x]})$, $x \in \mathbf{R}$, 而 $X_t = I(\phi_t)$, 其中 $\phi_t(x) = \mathrm{e}^{\mathrm{i}xt}$, $t \in \mathbf{R}$.

定理中的过程 Z 称为是平稳过程 X 的谱过程. 通俗地讲,平稳过程总是一个正交增量过程的 Fourier 变换.

习　题

1. 证明: 如果 X 是 $\sigma(X_t : t \in \mathsf{T})$ 可测的, 则存在可列集 $S \subset \mathsf{T}$, 使得 X 是 $\sigma(X_t : t \in S)$ 可测的.
2. 设 $(X_t : t \in \mathsf{T})$ 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上随机过程, X 是可积随机变量. 证明: 存在可列集 $S \subset \mathsf{T}$, 使得
$$\mathbb{E}(X|X_t : t \in \mathsf{T}) = \mathbb{E}(X|X_t : t \in S).$$
3. 设 K 是 $\mathsf{T} \times \mathsf{T}$ 上对称, 非负定实值函数. 证明: 存在一个 中心化的 Gauss 过程, 其协方差函数恰是 K.
4. 设 T 是个区间, $X = (X_t : t \in \mathsf{T})$ 是独立随机变量族, 且都服从以 p 为密度的分布, 证明: X 在任意点 $t \in \mathsf{T}$ 处都不是随机连续的. 问题说明, 至少在连续时间的情况下, 独立随机变量族不是一个好的随机过程.
5. 证明: X 是平稳独立增量过程当且仅当对任何 $s \geqslant 0$, 过程 $t \mapsto X_{t+s} - X_s$ 与 $\mathscr{F}_s$ 独立且与过程 $t \mapsto X_t - X_0$ 等价.
6. 设 T 是个区间. 证明: 实值随机过程 X 在 T 上随机连续当且仅当 $(t, s) \mapsto \mu_{(t,s)}$ 在 $\mathsf{T} \times \mathsf{T}$ 上以弱收敛的拓扑连续(弱连续).
7. 设 T 是个区间, 实值过程 X 随机连续. 证明: 如果 X 的几乎所有轨道在任意点 $t \in \mathsf{T}$ 处右极限存在, 则 X 有一个右连续修正.
8. 设 T 是个区间, X 是平方可积过程. 证明: X 在 T 上 L^2-连续当且仅当 $(t, s) \mapsto \mathbb{E}\overline{X_t}X_s$ 在 $\mathsf{T} \times \mathsf{T}$ 上连续.

9. 设 $f_1,\cdots,f_n$ 是区间 T 上的函数, $\xi_1,\cdots,\xi_n$ 是不相关随机变量, 方差分别为 $d_1,\cdots,d_n$. 令 $X_t:=\xi_1 f_1(t)+\cdots+\xi_n f_n(t)$. 计算随机过程 X 的协方差函数.
10. 如果过程 X,X' 是右连续的(或左连续), 那么它们互为修正蕴含着它们是不可区分的.
11. 设 A,η 是非负随机变量, θ 是独立于 A,η 且服从 $[0,2\pi)$ 上均匀分布的随机变量, 令 $X_t:=A\cos(\eta t+\theta)$, 证明: X 是一个广义平稳过程. 求 X 的谱过程.
12. 设 $\{X_n:\ n\in\mathbf{Z}\}$ 是独立随机变量列且 $P(X_n=-1)=P(X_n=1)=\dfrac{1}{2}$. 令

$$S_n:=\sum_{i=-\infty}^{n}\frac{1}{2^{|i|}}X_i,$$

显然 S_n 在点点收敛的意义下可以被定义. 验证 $\{S_n:\ n\in\mathbf{Z}\}$ 是一个正交增量过程, 其指标函数为

$$\lambda_S(n)=\sum_{i=-\infty}^{n}\frac{1}{4^{|i|}}.$$

§2.2 有限维分布族与相容定理

在上面, 我们定义了独立增量过程, 平稳过程, Gauss 过程等, 但定义本身不保证这样的随机过程存在, 这一节我们将主要讨论随机过程的存在性问题. 设 $(E,\mathscr{E})$ 是一个可测空间, 集合

$$\mathscr{D}=\{\mu_I: I\in\mathscr{I}_{\mathsf{T}}\}$$

称为是一个 E 上的有限维分布族, 如果对每个 $I\in\mathscr{I}_{\mathsf{T}}$, μ_I 是乘积空间 E^I 上的概率测度. E 上的有限维分布族 $\mathscr{D}=\{\mu_I: I\in\mathscr{I}_{\mathsf{T}}\}$ 称为是相容的, 如果它满足上节定理 2.1.1 的条件 (1), (2), 即横向相容与纵向相容. 从上一节的结论知道一个随机过程产生的有限维分布族总是相容的, 那么给定 E 上的一个相容的有限维分布族 $\mathscr{D}$, 是否存在一个概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 和其上的一个随机过程 X 使 X 的有限维分布族恰是给定的有限维分布族? 如果存在, 就说有限维分布族 $\mathscr{D}$ 可以实现, 而概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 和过程 X 是 $\mathscr{D}$ 的一个实现, 简单地重述上面的问题: 一个相容的有限维分布族是否一定可以实现?

首先让我们引入典则空间的概念, 它将在随机过程的构造理论中扮演重要的角色. 用 E^{T} 表示从 T 到 E 中的映射 $x=(x(t):t\in\mathsf{T})$ 全体组成的集合, E^{T} 中的元素有时也称为轨道, 而 E^{T} 通常称为轨道空间, 它实际上是 E 的 T 次自乘的乘积空间, 当 T 有限时, 就是通常的乘积空间. 对于 $x\in E^{\mathsf{T}}$, 令 $Z_t(x):=x(t)$, 它是 E^{T} 到 E 上的投影, 也称为坐标算子, 因为 $x(t)$ 也称为 x 在 t 处的坐标. 让 $\mathscr{E}^{\mathsf{T}}$ 是 E^{T} 上让所有投影 $\{Z_t:\ t\in\mathsf{T}\}$ 成为可测映射的最小 σ- 代数, 即

$$\mathscr{E}^{\mathsf{T}}:=\sigma(\{Z_t:t\in\mathsf{T}\}).$$

对 $I=(t_1,\cdots,t_n)\in\mathscr{I}_{\mathsf{T}}$, $x\in E^{\mathsf{T}}$, 令

$$\phi_I=\phi_{(t_1,\cdots,t_n)}(x):=(x(t_1),\cdots,x(t_n)),$$

则 ϕ_I 是 E^{T} 到 E^I 上的投影. 对任何 $H\in\mathscr{E}^I$,

$$\phi_I^{-1}(H)=\{x\in E^{\mathsf{T}}:\ (x(t_1),\cdots,x(t_n))\in H\}.$$

E^{T} 的形式如上的子集称为是 E^{T} 的一个柱集. 记 E^{T} 的所有柱集为

$$\mathscr{E}_0^{\mathsf{T}}:=\{\phi_I^{-1}(H):\ I\in\mathscr{I}_{\mathsf{T}},\ H\in\mathscr{E}^I\},$$

一般地 $\mathscr{E}_0^{\mathsf{T}}$ 不是 σ- 代数 (除非 T 是有限的).

引理 2.2.1　柱集的集合 $\mathscr{E}_0^{\mathsf{T}}$ 是一个代数且 $\mathscr{E}^{\mathsf{T}}=\sigma(\mathscr{E}_0^{\mathsf{T}})$.

证明　后一个结论是显然的. 现在证 $\mathscr{E}_0^{\mathsf{T}}$ 是一个代数. 由定义直接推出 $\varnothing,\Omega\in\mathscr{E}_0^{\mathsf{T}}$, 且 $\mathscr{E}_0^{\mathsf{T}}$ 对补集运算封闭. 另外容易看出, 对任何 $I=(t_1,\cdots,t_n)\in\mathscr{I}_{\mathsf{T}}$, $H\in\mathscr{E}^I$, $t\in\mathsf{T}$, 有

$$\phi_I^{-1}(H)=\phi_{I'}^{-1}(H'),$$

其中 $I':=(t_1,\cdots,t_{i-1},t,t_i,\cdots,t_n)\in\mathscr{I}_{\mathsf{T}}$. 而

$$H':=\{(x_1,\cdots,x_{n+1})\in E^{I'}:(x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_{n+1})\in H\},$$

那么对柱集 $\phi_{I_1}^{-1}(H_1)$, $\phi_{I_2}^{-1}(H_2)$, 存在 $I\in\mathscr{I}_{\mathsf{T}}$, $H_1',H_2'\in\mathscr{E}^I$, 使得

$$\phi_{I_1}^{-1}(H_1)=\phi_I^{-1}(H_1'),\ \phi_{I_2}^{-1}(H_2)=\phi_I^{-1}(H_2'),$$

则 $\phi_{I_1}^{-1}(H_1)\bigcup\phi_{I_2}^{-1}(H_2)=\phi_I^{-1}(H_1'\bigcup H_2')\in\mathscr{E}_0^{\mathsf{T}}$, 即 $\mathscr{E}_0^{\mathsf{T}}$ 对有限并封闭. $\square$

可测空间 $(E^{\mathsf{T}},\mathscr{E}^{\mathsf{T}})$ 也称为是典则空间. 典则空间上的坐标算子全体 $Z=(Z_t:t\in\mathsf{T})$ 是一个随机过程, 称为是典则过程.

设 $X=(X_t:t\in\mathsf{T})$ 是 $(\Omega,\mathscr{F},\mathbb{P})$ 上的 E-值随机过程, 则 Ω 到典则空间 E^{T} 有一个自然的映射, 它将 $\omega\in\Omega$ 映到 ω 在 E 上的轨迹: $t\mapsto X_t(\omega)$, $t\in\mathsf{T}$, 记此映射为 ξ, 即

$$\xi(\omega)(t)=X_t(\omega),\ \omega\in\Omega,t\in\mathsf{T}.$$

那么对任何 $I\in\mathscr{I}_{\mathsf{T}}$, $X_I=\phi_I\circ\xi$.下面的定理说明过程与典则过程的关系.

引理2.2.2 设 $\xi:\Omega\to E^{\mathsf{T}}$ 是如上定义的映射.

(1) ξ 是可测映射;

(2) X 与$(E^{\mathsf{T}},\mathscr{E}^{\mathsf{T}},\mathbb{P}\circ\xi^{-1})$ 上的典则过程 Z 是等价的, 其中 $\mathbb{P}\circ\xi^{-1}$ 是 $\mathbb{P}$ 在 ξ 下的像测度.

证明 取柱集 $\phi_I^{-1}(H)$, 则 $\xi^{-1}(\phi_I^{-1}(H))=X_I^{-1}(H)$. 因此 $\xi^{-1}(\mathscr{E}_0^{\mathsf{T}})\subset\mathscr{F}$, 从而 $\xi^{-1}(\mathscr{E}^{\mathsf{T}})\subset\mathscr{F}$, 证明了 (1). 为证 (2), 任取 $I\in\mathscr{I}_{\mathsf{T}}$, X 的有限维分布 $\mu_I=\mathbb{P}\circ X_I^{-1}=(\mathbb{P}\circ\xi^{-1})\circ\phi_I^{-1}$, 最后的概率是 Z 的有限维分布. □

说一个有限维分布族可以在典则空间上实现, 如果典则空间上有一个概率 $\mathbb{P}$ 使得典则过程的有限维分布族恰是给定的有限维分布族. 下面引理实际上是引理2.2.2的另一种叙述方式.

引理 2.2.3 一个有限维分布族有一个实现当且仅当它可以在典则空间上实现.

下面我们将证明本节的主要结果, 它通常也称为 Kolmogorov相容定理.

定理 2.2.1 (Kolmogorov) 设 E 是完备可分度量空间, $\mathscr{E}$ 是对应的 Borel σ- 代数, 则 $(E,\mathscr{E})$ 上的任何相容的有限维分布族一定有一个实现.

证明 设 $\mathscr{D}=\{\mu_I:\ I\in\mathscr{I}_{\mathsf{T}}\}$ 是 $(E,\mathscr{E})$ 上的一个相容的有限维分布族. 让我们先在柱集类 $\mathscr{E}_0^{\mathsf{T}}$ 上构造一个集函数 $\mathbb{P}$ 如下:

$$\mathbb{P}(\phi_I^{-1}(H)):=\mu_I(H),\ I\in\mathscr{I}_{\mathsf{T}},\ H\in\mathscr{E}^I.$$

一个柱集可以有不同的表示, 但由分布族的相容性, 上述定义不会产生歧义. 例如 $\phi_{(t_1,t_2)}^{-1}(A_1\times A_2)=\phi_{(t_2,t_1)}^{-1}(A_2\times A_1)$, 但由 $\mathscr{D}$ 的横向相容性, $\mu_{(t_1,t_2)}(A_1\times A_2)=$

$\mu_{(t_2,t_1)}(A_2 \times A_1)$; 又 $\phi^{-1}_{(t_1,t_2)}(A_1 \times E) = \phi^{-1}_{(t_1)}(A_1)$, 但同样 $\mathscr{D}$ 的纵向相容性推出 $\mu_{(t_1,t_2)}(A_1 \times E) = \mu_{(t_1)}(A_1)$. (读者不妨自己写出一个严格的证明.) 因为 $\mathscr{E}_0^{\mathsf{T}}$ 是一个代数, 由测度扩张定理, 只要能证明 $\mathbb{P}$ 是 $\mathscr{E}_0^{\mathsf{T}}$ 上的一个预概率测度, 它便可以扩张到 $\mathscr{E}^{\mathsf{T}}$ 上成为一个概率测度, 而且由上面的定义容易看出在 $\mathbb{P}$ 下典则过程的有限维分布族恰是 $\mathscr{D}$.

容易验证 $\mathbb{P}$ 在 $(E^{\mathsf{T}}, \mathscr{E}_0^{\mathsf{T}})$ 上有下列性质:

(1) $\mathbb{P}(E^{\mathsf{T}}) = 1$, $\mathbb{P}(\varnothing) = 0$;

(2) 如果 $A \in \mathscr{E}_0^{\mathsf{T}}$, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$;

(3) (有限可加性) 如果 $A, B \in \mathscr{E}_0^{\mathsf{T}}$ 且 $A \cap B = \varnothing$, 则

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

我们仅需验证 $\mathbb{P}$ 是上连续的, 即取一列单调下降且交为空集的柱集 $\{A_n\}$ 必有 $\mathbb{P}(A_n) \downarrow 0$. 假设不然, 存在 $\epsilon > 0$, 使 $\mathbb{P}(A_n) > \epsilon$. 对于柱集列, 我们总可以取一个时间列 $\{t_n\} \subset \mathsf{T}$ 及 $H_n \in \mathscr{E}^n$, 使得 $A_n = \phi^{-1}_{(t_1,\cdots,t_n)}(H_n)$, 那么 $\mu_{(t_1,\cdots,t_n)}(H_n) = \mathbb{P}(A_n) > \epsilon$. 因 E 从而 E^n 是完备可分度量空间, 任何其上的有限测度都是正则的(参考[28]), 即有紧集 $K_n \subset H_n$, 使

$$\mu_{(t_1,\cdots,t_n)}(H_n \backslash K_n) < \frac{\epsilon}{2^n}.$$

再令 $B_n := \phi^{-1}_{(t_1,\cdots,t_n)}(K_n)$, 则

$$\mathbb{P}(A_n \backslash B_n) = \mu_{(t_1,\cdots,t_n)}(H_n \backslash K_n) < \frac{\epsilon}{2^n}.$$

记 $C_n := \bigcap_{k \leqslant n} B_k$, 那么

$$\begin{aligned}\mathbb{P}(A_n - C_n) &= \mathbb{P}\Big(\bigcup_{k \leqslant n} (A_n \backslash B_k)\Big) \\ &\leqslant \mathbb{P}\Big(\bigcup_{k \leqslant n} (A_k \backslash B_k)\Big) \\ &\leqslant \sum_{k \leqslant n} \mathbb{P}(A_k \backslash B_k) < \epsilon.\end{aligned}$$

故 $\mathbb{P}(C_n) > \mathbb{P}(A_n) - \epsilon > 0$, C_n 自然是非空的, 取 $x^{(n)} \in C_n$. 因 $\{C_n\}$ 是单调下降的, 故对 $l \leqslant n$, $x^{(n)} \in C_l \subset B_l$, 即

$$(x^{(n)}(t_1), x^{(n)}(t_2), \cdots, x^{(n)}(t_l)) \in K_l.$$

由于 K_l 紧, 任意固定 l, 点列 $\{x^{(n)}(t_l):\ n \geqslant 1\}$ 有收敛子列, 由对角线法, 存在一个自然数子列 $\{n_i\}$, 使对任何 l, $\{x^{(n_i)}(t_l)\}_{i\geqslant 1}$ 收敛, 令 $x_l := \lim_i x^{(n_i)}(t_l)$. 取 $x \in E^{\mathsf{T}}$, 使 $x(t_l) = x_l$, 则 $(x(t_1), \cdots, x(t_l)) \in K_l$. 因此对任何 $l \geqslant 1$, $x \in B_l \subset A_l$, 故 $\bigcap_{l\geqslant 1} A_l$ 非空, 这导致矛盾. □

注意到在上面定理中 E 要求是一个完备可分的度量空间, 这不是必需的, 对状态空间的拓扑要求可以减弱, 但是, E 仅是一般的可测空间是不够的. 定理中构造的概率 $\mathbb{P}$ 由有限维分布唯一确定, 也称为是给定的有限维分布族 $\mathscr{D}$ 的极限.

作为相容性定理的应用, 我们来证明无穷维乘积概率空间的存在性. 设 $\{\mu_t : t \in \mathsf{T}\}$ 是 $(E, \mathscr{E})$ 上的概率测度族, 对 $I = (t_1, \cdots, t_n) \in \mathsf{T}$, $B_1, \cdots, B_n \in \mathscr{E}$, 令

$$\mu_{(t_1,\cdots,t_n)}(B_1 \times \cdots \times B_n) := \prod_{i=1}^{n} \mu_{t_i}(B_i).$$

容易验证 $\{\mu_I : I \in \mathscr{I}_{\mathsf{T}}\}$ 是一个相容的有限维分布族. 故我们有下列推论.

推论 2.2.1 设 $\{\mu_t : t \in \mathsf{T}\}$ 是完备可分度量空间 E 上的概率分布族, 则存在概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 及其上的随机过程 $X = (X_t : t \in \mathsf{T})$, 使得:

(1) $\{X_t : t \in \mathsf{T}\}$ 是独立的;

(2) 对任何 $t \in \mathsf{T}$,X_t 的分布是 μ_t.

特别地, 如果 μ 是 E 上的概率, 则存在分布同为 μ 的独立随机变量序列.

我们将进一步讨论 $\mathscr{E}^{\mathsf{T}}$ 的性质. 首先, 说 E^{T} 的一个子集 A 是可列决定的, 如果存在一个可列集 $S \subset \mathsf{T}$, 使得对 $x \in A$, $y \in E^{\mathsf{T}}$, 如果对任何 $t \in S$, 有 $x(t) = y(t)$, 则 $y \in A$. 也说 A 可由 S 决定. 令 $\mathscr{A}$ 为 E^{T} 的可列决定的子集全体. 它是 E^{T} 上的 σ- 代数, 事实上, 我们仅需要验证 $\mathscr{A}$ 关于补运算及可列并运算封闭, 设 $A \in \mathscr{A}$ 且 A 可由 S 决定, 则 A^c 也可由 S 决定, 因为如果 $x \in A^c$, $y \in E^{\mathsf{T}}$ 且两者在 S 上重合, 则 $y \in A$ 蕴含着 $x \in A$, 故 $y \in A^c$. 再设 $A_n \in \mathscr{A}$ 且 A_n 可由 S_n 决定, 则显然 $\bigcup_{n\geqslant 1} A_n$ 可由可列集 $\bigcup_{n\geqslant 1} S_n$ 决定, 即 $\bigcup_{n\geqslant 1} A_n \in \mathscr{A}$. 另外容易看出任何柱集

是可列决定的, 即$\mathscr{E}_0^{\mathsf{T}} \subset \mathscr{A}$, 因此 $\mathscr{E}^{\mathsf{T}} \subset \mathscr{A}$.

定理 2.2.2 $\mathscr{E}^{\mathsf{T}}$ 中的集合是可列决定的.

上述定理说明 $\mathscr{E}^{\mathsf{T}}$ 中的集合结构很简单, 一般来说, 改变集 $A \in \mathscr{E}^{\mathsf{T}}$ 中一个轨道 x 的某个点的值不会将 x 移出 A. 换句话说, 如果改变 E^{T} 的子集 A 的一个轨道 x 在某点处的值会导致新的轨道不再属于 A, 则 A 不可能在 $\mathscr{E}^{\mathsf{T}}$ 中. 因此许多重要的通常是有某种正则性的轨道集合不是 $\mathscr{E}^{\mathsf{T}}$ 可测集, 如连续轨道的集合, 右连续且存在左极限的轨道集合等, 因为直观地, 改变一条连续轨道在某点处的值会导致轨道不再连续. 让我们详细地进行解释.

设 E 是一个至少含有两个点的 Hausdorff 拓扑空间, $\mathsf{T} = \mathbf{R}^+$, 则 T 到 E 的连续映射全体 $C(\mathsf{T}; E)$ 不是可列决定的, 即 $C(\mathsf{T}; E) \notin \mathscr{E}^{\mathsf{T}}$. 事实上, 假设决定 $C(\mathsf{T}; E)$ 的一个可列集 S 存在, 任取 $\{a, b\} \subset E$, $a \neq b$, 令

$$x(t) = a,\ t \geqslant 0;\ y(t) = \begin{cases} a, & t \in S; \\ b, & t \notin S, \end{cases}$$

则 x 与 y 在 S 上重合, 但 $x \in C(\mathsf{T}; E)$, 而 $y \notin C(\mathsf{T}; E)$. 矛盾. 同样可以证明 $C(\mathsf{T}; E)$ 不包含任何可测集.

因此仅讨论 $\mathscr{E}^{\mathsf{T}}$ 中的集合显然是不够的, 故我们下面将引入本质轨道空间的概念. 设 $X = (X_t : t \in \mathsf{T})$ 是 $(\Omega, \mathscr{F}, \mathbb{P})$ 上以 $(E, \mathscr{E})$ 为状态空间的随机过程, ξ 是如前定义的 Ω 到 E^{T} 的映射. E^{T} 的一个子集 $\tilde{\Omega}$ 称为是 X 的一个轨道空间, 如果 $\tilde{\Omega} \supset \xi(\Omega)$. 注意, 这里 $\tilde{\Omega}$ 并不被要求是可测的, 即不一定要属于 $\mathscr{E}^{\mathsf{T}}$.

定义 2.2.1 设 $\mathscr{D}$ 是 $(E, \mathscr{E})$ 上有限维分布族, E^{T} 的一个子集 $\tilde{\Omega}$ 称为关于$\mathscr{D}$是本质的, 如果存在以 $(E, \mathscr{E})$ 为状态空间以 $\mathscr{D}$ 为有限维分布族的一个随机过程 X, 使得 $\tilde{\Omega}$ 是 X 的一个轨道空间.

定理 2.2.3 设 E 是完备可分度量空间, $\hat{\mathbb{P}}$ 是相容的有限维分布族 $\mathscr{D}$ 的极限, 则 $\tilde{\Omega} \subset E^{\mathsf{T}}$ 是本质的当且仅当 $\hat{\mathbb{P}}^*(\tilde{\Omega}) = 1$, 其中 $\hat{\mathbb{P}}^*$ 是 $\hat{\mathbb{P}}$ 的外测度.

证明 首先设 $\tilde{\Omega} \in E^{\mathsf{T}}$ 关于 $\mathscr{D}$ 是本质的, 由引理 2.2.2, 因为典则空间上的测度 $\hat{\mathbb{P}}$ 与 $\mathbb{P} \circ \xi^{-1}$ 有同样的有限维分布族, 因此 $\hat{\mathbb{P}} = \mathbb{P} \circ \xi^{-1}$. 对任何 $A \in \mathscr{E}^{\mathsf{T}}$ 且 $\tilde{\Omega} \subset A$, 则

$\Omega \subset \xi^{-1}(\tilde{\Omega}) \subset \xi^{-1}(A)$, 故 $\hat{\mathbb{P}}(A) \geqslant \mathbb{P}(\Omega) = 1$. 因此 $\hat{\mathbb{P}}^*(\tilde{\Omega}) = 1$. 反之, 如果 $\hat{\mathbb{P}}^*(\tilde{\Omega}) = 1$, 那么令

$$\tilde{\mathscr{F}} := \tilde{\Omega} \cap \mathscr{E}^{\mathsf{T}} = \{\tilde{\Omega} \cap A : A \in \mathscr{E}^{\mathsf{T}}\}.$$

当然 $(\tilde{\Omega}, \tilde{\mathscr{F}})$ 是可测空间, 定义 $\tilde{\mathscr{F}}$ 上的集函数:

$$\tilde{\mathbb{P}}(\tilde{\Omega} \cap A) := \hat{\mathbb{P}}(A), \ A \in \mathscr{E}^{\mathsf{T}}.$$

为使定义无歧义, 我们要验证对 $A, B \in \mathscr{E}^{\mathsf{T}}$, 如果 $\tilde{\Omega} \cap A = \tilde{\Omega} \cap B$, 则 $\hat{\mathbb{P}}(A) = \hat{\mathbb{P}}(B)$. 事实上, 这时 $A \triangle B \subset \tilde{\Omega}^c$, 故 $\tilde{\Omega} \subset (A \triangle B)^c$, 因此 $\hat{\mathbb{P}}(A \triangle B) = 0$, 即有 $\hat{\mathbb{P}}(A) = \hat{\mathbb{P}}(B)$. 然后对 $\tilde{\omega} \in \tilde{\Omega}$ 和 $t \in \mathsf{T}$, 令 $\tilde{X}_t(\tilde{\omega}) = Z_t(\tilde{\omega})$. 容易验证 $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}})$ 上的随机过程 $\tilde{X} = (\tilde{X}_t : t \in \mathsf{T})$ 的有限维分布族是 $\mathscr{D}$, 因此 $\tilde{\Omega}$ 是本质的. □

这个定理说明, $\mathscr{E}^{\mathsf{T}}$ 可测集太少不会引起本质的问题, 我们总可以在外测度为 1 的集合上重新定义概率空间与随机过程.

习 题

1. 设 $\mathsf{T} = [0,1]$, $\mathscr{B}$ 是 $\mathbf{R}$ 上 Borel 代数. 验证: 线性函数的集合, 多项式的集合, 递增函数的集合, 在某一固定点连续的集合, Borel 可测函数的集合都不属于 $\mathscr{B}^{\mathsf{T}}$.
2. 设取值 0, 1 的随机序列 $(X_n : n \geqslant 1)$ 满足:

$$\lim_n \mathbb{P}(\bigcup_{k \geqslant n} \{X_k = 1\}) = 1.$$

   而 μ 是随机变量 $\sum\limits_{n \geqslant 1} \dfrac{X_n}{2^n}$ 的分布. 对 $x \in (0,1]$, 让 $\xi_n(x)$ 是 x 的二进表示的第 n 位小数. 证明: 在 $([0,1], \mu)$ 上的序列 $\{\xi_n\}$ 与 $\{X_n\}$ 等价.
3. 设 $\{\mathscr{F}_n : n \geqslant 1\}$ 是 Ω 上一个递增 σ- 代数列, $\mathbb{P}_n$ 是 $(\Omega, \mathscr{F}_n)$ 上的概率且满足 $\mathbb{P}_{n+1}|_{\mathscr{F}_n} = \mathbb{P}_n$, 令 $\mathscr{F} := \sigma(\mathscr{F}_n : n \geqslant 1)$. 问: $(\Omega, \mathscr{F})$ 上是否存在概率 $\mathbb{P}$, 满足对任何 n, 有 $\mathbb{P}|_{\mathscr{F}_n} = \mathbb{P}_n$?
4. 证明: 区间 $[0,1]$ 上不可能有不可数个非常数的独立随机变量.

§2.3　Markov过程与转移半群

首先让我们引入核及核诱导的算子等概念.

定义 2.3.1　设 $(\Omega,\mathscr{F})$ 和 $(\Omega',\mathscr{F}')$ 是两个可测空间, 映射 $K:\Omega\times\mathscr{F}'\longrightarrow[0,\infty]$ 称为是从 $(\Omega,\mathscr{F})$ 到 $(\Omega',\mathscr{F}')$ 的核, 如果它满足如下条件:

(1) 对固定的 $A'\in\mathscr{F}'$, $K(\cdot,A')$ 是 $(\Omega,\mathscr{F})$上可测函数;

(2) 对固定的 $\omega\in\Omega$, $K(\omega,\cdot)$ 是$(\Omega',\mathscr{F}')$ 上的测度.

当 $\Omega=\Omega'$, $\mathscr{F}=\mathscr{F}'$ 时, 称 K 是 $(\Omega,\mathscr{F})$ 上的核. 这时, 如果 $K(\cdot,\Omega)=1$, 称 K 是 Markov 核, 而若 $K(\cdot,\Omega)\leqslant 1$, 称 K 是子 Markov 核.

任何子 Markov 核总可以适当地修改而成为 Markov 核, 事实上, 设 K 是 $(\Omega,\mathscr{F})$ 上的子 Markov 核, 取 $\Delta\notin\Omega$, 定义 $\Omega_\Delta:=\Omega\cup\{\Delta\}$, $\mathscr{F}_\Delta:=\sigma(\mathscr{F}\cup\{\{\Delta\}\})$. 对 $\omega\in\Omega_\Delta$, $A\in\mathscr{F}_\Delta$, 定义

$$K'(\omega,A):=\begin{cases}K(\omega,A), & \omega\in\Omega,\ A\in\mathscr{F},\\ 1-K(\omega,\Omega), & \omega\in\Omega,\ A=\{\Delta\},\\ 1_A(\Delta), & \omega=\Delta,\ A\in\mathscr{F}_\Delta.\end{cases}$$

容易验证 K' 是 $(\Omega_\Delta,\mathscr{F}_\Delta)$ 上的 Markov 核.

设 K 是 $(\Omega,\mathscr{F})$ 到 $(\Omega',\mathscr{F}')$ 的核, 任取 Ω' 上的非负可测函数 f, 记

$$Kf(\omega):=\int_{\Omega'}K(\omega,\mathrm{d}\omega')f(\omega'),\ \omega\in\Omega,$$

则 Kf 是 Ω上的非负可测函数, 即 K 诱导一个将 Ω'上的非负可测函数映至 Ω上的非负可测函数的算子, 称为是拉回算子. 反过来, 如果 μ 是 $(\Omega,\mathscr{F})$ 上的一个测度, 记

$$\mu K(A'):=\int_{\Omega}\mu(\mathrm{d}\omega)K(\omega,A'),\ A'\in\mathscr{F}',$$

则 μK 是 $(\Omega;\mathscr{F}')$ 上的一个测度, 即 K 诱导一个将 $(\Omega,\mathscr{F})$ 上的测度映至 $(\Omega',\mathscr{F}')$ 上的测度的算子, 称为推前算子. 显然, 当 K 是 Markov 核时, 拉回算子将有界可测

函数映为有界可测函数, 而推前算子将概率测度映为概率测度.

设 K_1 和 K_2 分别是 $(\Omega_1,\mathscr{F}_1)$ 到 $(\Omega_2,\mathscr{F}_2)$ 和 $(\Omega_2,\mathscr{F}_2)$ 到 $(\Omega_3,\mathscr{F}_3)$ 的核, 对 $\omega_1\in\Omega_1$, $A_3\in\mathscr{F}_3$, 定义

$$\begin{aligned}K_1K_2(\omega_1,A_3)&:=K_1(K_2(\cdot,A_3))(\omega_1)=(K_1(\omega_1,\cdot)K_2)(A_3)\\&=\int K_1(\omega_1,\mathrm{d}\omega_2)K_2(\omega_2,A_3),\end{aligned}$$

则 K_1K_2 是一个 $(\Omega_1,\mathscr{F}_1)$ 到 $(\Omega_3,\mathscr{F}_3)$ 的核.

现在我们将引入转移半群. 设 $(E,\mathscr{E})$ 是一个可测空间, T 是非负整数或非负实数集.

定义 2.3.2 $(E,\mathscr{E})$ 上的一族(子) Markov 核 $(P_t:\ t\in\mathsf{T})$ 称为是转移半群, 如果对任意 $s,t\in\mathsf{T}$, 有

$$P_sP_t=P_{t+s}.$$

当 P_0 是恒等算子时, 称转移半群是正规的.

注 如同前面的核一样, 一个子 Markov 核组成的转移半群总可以通过引入一个新点重定义成为 Markov 核组成的转移半群, 因此下面我们说到转移半群是指由 Markov 核组成的. 另外公式 $P_{t+s}=P_sP_t$ 等价于

$$P_{t+s}(x,B)=\int_E P_s(x,\mathrm{d}y)P_t(y,B),\ x\in E,B\in\mathscr{E},$$

称为是 Chapman-Kolmogorov 方程.

设 $(P_t:t\in\mathsf{T})$ 是可测空间 $(E,\mathscr{E})$ 上的转移半群, 对任何 $n\geqslant 1$, $t_1,\cdots,t_n\in\mathsf{T}$, $0\leqslant t_1<\cdots<t_n$ 及 $A\in\mathscr{E}^n$, $x\in E$, 定义

$$\mathbb{P}_I(x,A):=\int_A P_{t_1}(x,\mathrm{d}x_1)P_{t_2-t_1}(x_1,\mathrm{d}x_2)\cdots P_{t_n-t_{n-1}}(x_{n-1},\mathrm{d}x_n),$$

其中 I 表示 $(t_1,\cdots,t_n)$, 则 $\mathbb{P}_I$ 是 $(E,\mathscr{E})$ 到 $(E^I,\mathscr{E}^I)$ 的核. 任取 E 上的概率测度 μ, 再定义

$$\mu_I:=\mu\mathbb{P}_I=\int_{x\in E}\mu(\mathrm{d}x)\mathbb{P}_I(x,\cdot),$$

注意 $\mu P_{(0)} = \mu P_0$.

定理 2.3.1　对任何 $(E, \mathscr{E})$ 上给定的概率测度 μ, 上面定义的概率测度族 $\mathscr{D}_\mu = \{\mu_I : I \in \mathscr{I}_{\mathsf{T}}\}$ 是 $(E, \mathscr{E})$ 上的相容的有限维分布族.

证明　$\mathscr{D}_\mu$ 的相容性由 Fubini 定理与 Chapman-Kolmogorov 方程容易推出. 留给读者作为练习. □

现在我们介绍所谓的 Markov 性, 先引入流的概念, 它也是描述随机过程的一个方便工具. 设 $(\Omega, \mathscr{F}, \mathbb{P})$ 是一个概率空间, T 是时间集. 如果 $\mathscr{F}$ 的一个子 σ- 代数族 $\{\mathscr{F}_t : t \in \mathsf{T}\}$ 满足对任何 $s, t \in \mathsf{T}, s < t$, 有 $\mathscr{F}_s \subset \mathscr{F}_t$, 称它是 Ω 上的 σ- 代数流或简称流. 随机过程 $X = (X_t : t \in \mathsf{T})$ 称为是 $(\mathscr{F}_t)$ 适应的, 如果对任何 $t \in \mathsf{T}$, X_t 是 $\mathscr{F}_t$ 可测的. 给定一个随机过程 $X = (X_t : t \in \mathsf{T})$, 定义

$$\mathscr{F}_t := \sigma(\{X_s : s \leqslant t\}),\ t \in \mathsf{T},$$

即 $\mathscr{F}_t$ 是时刻 t 之前的过程所生成的 σ- 代数. 它显然构成一个 Ω 上的流, 称之为 X 的自然流, X 关于其自然流总是适应的, 必要时注明为 $(\mathscr{F}_t^X)$. 它是 X 所适应的流 中最小的, 它通常解释为 t 时刻前过程的信息.

设 $(\mathscr{F}_t)$ 是一个流, 对任何 $t \in \mathsf{T}$, 定义 $\mathscr{F}_{t+} := \bigcap_{s>t} \mathscr{F}_s$, 显然 $(\mathscr{F}_{t+})_{t \in \mathsf{T}}$ 也是一个流. 如果对任何 $t \in \mathsf{T}$, $\mathscr{F}_{t+} = \mathscr{F}_t$, 称流 $(\mathscr{F}_t)$ 是右连续的. 当然, 流 $(\mathscr{F}_{t+})$ 是右连续的且一个适应于 $(\mathscr{F}_t)$ 的过程也一定适应于 $(\mathscr{F}_{t+})$. 如果 $(\Omega, \mathscr{F}, \mathbb{P})$ 是完备的且 $\mathscr{F}_0$ 含有所有 $\mathbb{P}$-零测集, 则称流 $(\mathscr{F}_t)$ 是完备的.

给定概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 和一个流 $(\mathscr{F}_t)$, 这时我们简单地说带流概率空间 $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$. X 是以 $(E, \mathscr{E})$ 为状态空间的 $(\mathscr{F}_t)$-适应的随机过程. 称随机过程 X 相对于流 $(\mathscr{F}_t)$ 具有 Markov 性(或 $(\mathscr{F}_t)$-Markov 过程), 如果对 $s, t \in \mathsf{T}$, $s < t$ 及 $B \in \mathscr{E}$, 有

$$\mathbb{P}(X_t \in B | \mathscr{F}_s) = \mathbb{P}(X_t \in B | X_s). \tag{M1}$$

这时 X 关于其自然流也一定有 Markov 性. 如果不特别地提到流, 我们总认为 X 关于其自然流有 Markov 性. 由 Dynkin 的方法, 不难证明(作为习题)过程 X 有 Markov 性当且仅当对任何 $s_1, \cdots, s_n, t \in \mathsf{T}$ 且 $0 \leqslant s_1 < \cdots < s_n < t$, 有

$$\mathbb{P}(X_t \in B | X_{s_1}, \cdots, X_{s_n}) = \mathbb{P}(X_t \in B | X_{s_n}). \tag{M2}$$

此性质解释为在已知现在位置的条件下, 将来的位置与过去是独立的. 也就是说, 分

析一个 Markov 过程的过去的数据无助于对将来的预测. 对 Markov 性, 在习题中有更直观的刻画. 取例 2.1.1 来说, 如果已知第 n 次赌博后甲所持有的赌资, 那么 $n+k$ 次赌博后, 甲所持有的赌资与 n 以前他的输赢无关. 这是一个直观的但却常被人们忽略或不愿承认的事实. 一些赌徒会固执地认为在连输几局后, 下一次获胜的概率会增大.

例 2.3.1 我们来证明独立增量过程(如例 2.1.6)总是 Markov 过程. 带流概率空间 $(\Omega,\mathscr{F},(\mathscr{F}_t),\mathbb{P})$ 上的取值在 $\mathbf{R}^d$ 的适应过程 $X=(X_t:t\geqslant 0)$ 称为是独立增量的, 如果对任何 $t>s\geqslant 0$, X_t-X_s 与 $\mathscr{F}_s$ 独立. 这时, 任取 $B\in\mathscr{B}(\mathbf{R}^d)$, 令 $A:=\{(x,y):x,y\in\mathbf{R}^d,x+y\in B\}$. 因为 X_t-X_s 独立于 $\mathscr{F}_s$, 而 X_s 是 $\mathscr{F}_s$ 可测, 故

$$\begin{aligned}\mathbb{P}(X_t\in B|\mathscr{F}_s)&=\mathbb{P}((X_t-X_s,X_s)\in A|\mathscr{F}_s)\\&=\mathbb{P}((X_t-X_s,y)\in A)\big|_{y=X_s}=\mathbb{P}(X_t\in B|X_s).\end{aligned}$$

证明了 Markov 性. ∎

下面定理说明转移半群生成的有限维分布族的实现一定是 Markov 过程.

定理 2.3.2 设 $\mathscr{D}_\mu$ 是定理 2.3.1 中所定义的有限维分布族, 概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 上的 E-值随机过程 X 是它的一个实现, 则 X 是 Markov 过程. 更进一步, $(M1)$ 等价于对 $s,t\in\mathsf{T},s<t$ 且 $B\in\mathscr{E}$,

$$\mathbb{P}(X_t\in B|\mathscr{F}_s)=P_{t-s}(X_s,B).\tag{$M3$}$$

证明 我们只需验证对 $\{s_1,\cdots,s_n,t\}\subset\mathsf{T}$ 且 $0\leqslant s_1<\cdots<s_n=s<t$, 有

$$\mathbb{P}(X_t\in B|X_{s_1},\cdots,X_{s_n})=P_{t-s}(X_s,B),\ \text{a.s.}.$$

取 $I=(s_1,\cdots,s_n)$, $A\in\mathscr{E}^I$,

$$\begin{aligned}&\mathbb{E}(1_{\{X_t\in B\}};\{X_I\in A\})\\&=\mathbb{P}((X_{s_1},\cdots,X_{s_n})\in A,X_t\in B)\\&=\int_{x\in E,(x_1,\cdots,x_n)\in A}\mu(\mathrm{d}x)P_{s_1}(x,\mathrm{d}x_1)\cdots P_{s_n-s_{n-1}}(x_{n-1},\mathrm{d}x_n)P_{t-s_n}(x_n,B)\end{aligned}$$

$$= \int_{E^I} \mu_I(\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n)1_A(x_1,\cdots,x_n)P_{t-s_n}(x_n,B)$$

$$= \mathbb{E}(P_{t-s}(X_s,B);\{(X_{s_1},\cdots,X_{s_n})\in A\}).$$

完成证明. □

如果 (P_t) 是正规的, 那么 $(M3)$ 当 $s=t$ 时也成立. 这时, X_0 的分布也恰好是 $\mu P_0=\mu$, 以上定理中的过程 X 称为是具有转移概率 (P_t) 与初始分布 μ 的 Markov 过程.

现在设 $(E,\mathscr{E})$ 是完备可分度量空间及其 Borel 代数, $(\Omega,\mathscr{F})$ 是典则空间, $X=(X_t:t\in\mathsf{T})$ 是典则过程, $(P_t)_{t\in\mathsf{T}}$ 是 $(E,\mathscr{E})$ 上的转移半群, 则由 Kolmogorov 定理, 对任何 $x\in E$, 存在 $(\Omega,\mathscr{F})$ 上的概率测度 $\mathbb{P}^x$, 使得典则过程 $X=(X_t)$ 的有限维分布族为 $\{\mathbb{P}_I(x,\cdot):I\in\mathscr{I}_\mathsf{T}\}$, 即

$$\mathbb{P}^x((X_{t_1},\cdots,X_{t_n})\in A)=P_{(t_1,\cdots,t_n)}(x,A)$$
$$=\int_A P_{t_1}(x,\mathrm{d}x_1)P_{t_2-t_1}(x_1,\mathrm{d}x_2)\cdots P_{t_n-t_{n-1}}(x_{n-1},\mathrm{d}x_n),\ A\in\mathscr{E}^n.$$

或者说 $(X_{t_1},\cdots,X_{t_n})$ 在概率 $\mathbb{P}^x$ 下的联合分布为

$$P_{t_1}(x,\mathrm{d}x_1)P_{t_2-t_1}(x_1,\mathrm{d}x_2)\cdots P_{t_n-t_{n-1}}(x_{n-1},\mathrm{d}x_n).$$

由核的定义, 上式右边是关于 x 的可测函数, 因此上式左边也是. 因此对任何 $G\in\mathscr{F}$, $x\mapsto\mathbb{P}^x(G)$ 是 E 上可测函数, 即 $(x,G)\mapsto\mathbb{P}^x(G)$ 是 $(E,\mathscr{E})$ 到 $(\Omega,\mathscr{F})$ 的 Markov 核. 这时上面定理 2.3.2 的结论可写为

$$\mathbb{P}^x(X_{t+s}\in B|\mathscr{F}_t)=\mathbb{P}^{X_t}(X_s\in B),\ \text{a.s.}, \tag{M4}$$

其中 $t,s\in\mathsf{T}$, $s,t>0$, $x\in E$ 及 $B\in\mathscr{E}$. 如果 (P_t) 是正规的, 则上式对 $s,t\in\mathsf{T}$ 都是对的, 且 X_0 的分布是 x 的单点测度.

典则空间上有一个自然的推移算子族, 对 $\omega\in\Omega$, $t\in\mathsf{T}$, 定义

$$\theta_t\omega(s)=\omega(t+s),\ s\in\mathsf{T},$$

则 θ_t 是 $(\Omega,\mathscr{F})$ 上可测映射, 满足 $X_s\circ\theta_t=X_{s+t}$. 那么上面的 Markov 性等价于说, 对 $(\Omega,\mathscr{F})$ 上有界或非负随机变量 Y 与 $x\in E$, $t\in\mathsf{T}$, 有

$$\mathbb{E}^x(Y\circ\theta_t|\mathscr{F}_t)=\mathbb{E}^{X_t}(Y),\ \text{a.s.}. \tag{$M4'$}$$

事实上, 显然 $(M4')$ 可推出 $(M4)$, 反之对 $0<t_1<\cdots<t_n$, 及 Borel 集 $A_1,\cdots,A_n$, 取 $t=t_i$,

$$\begin{aligned}
&\mathbb{P}^x(X_{t_1}\in A_1,\cdots,X_{t_n}\in A_n)\\
&\quad=\int_{x_1\in A_1,\cdots,x_i\in A_i}P_{t_1}(x,\mathrm{d}x_1)P_{t_2-t_1}(x_1,\mathrm{d}x_2)\cdots P_{t-t_{i-1}}(x_{i-1},\mathrm{d}x_i)\\
&\qquad\int_{x_{i+1}\in A_{i+1},\cdots,x_n\in A_n}P_{t_{i+1}-t}(x_i,\mathrm{d}x_{i+1})\cdots P_{t_n-t_{n-1}}(x_{n-1},\mathrm{d}x_n)\\
&\quad=\mathbb{E}^x(\mathbb{P}^{X_t(\cdot)}(X_{t_{i+1}-t}\in A_{i+1},\cdots,X_{t_n-t}\in A_n);X_t\in A_i,\cdots,X_{t_1}\in A_1).
\end{aligned}$$

这蕴含着 $(M4')$ 对示性函数 $Y=1_B$ 成立, 其中 $B=\{X_{t_{i+1}-t}\in A_{i+1},\cdots,X_{t_n-t}\in A_n\}$ 是柱集. 因为柱集是 π-类, 故由 Dynkin 定理, $(M4')$ 对 $Y=1_B$, $B\in\mathscr{F}$ 成立, 然后由单调收敛定理 $(M4')$ 对非负随机变量也成立.

定义 2.3.3 设 $(E,\mathscr{E})$ 是可测空间, 一个六元组

$$X=(\Omega,\mathscr{F},\mathscr{F}_t,X_t,\theta_t,\mathbb{P}^x:t\in\mathsf{T},x\in E)$$

称为是一个以 $(E,\mathscr{E})$ 为状态空间的时齐 $(\mathscr{F}_t)$-Markov 过程, 如果

(1) 对任何 $x\in E$, $\mathbb{P}^x$ 是 $(\Omega,\mathscr{F})$ 上的概率, $(X_t)_{t\in\mathsf{T}}$ 是以 E 为状态空间的适应随机过程;

(2) 对 $A\in\mathscr{F}$, $x\mapsto\mathbb{P}^x(A)$ 在 $(E,\mathscr{E})$ 上可测;

(3) 对 $t\in\mathsf{T}$, $\theta_t:\Omega\to\Omega$ 可测且对任何 $s\in\mathsf{T}$ 满足 $X_s\circ\theta_t=X_{s+t}$;

(4) 对 $(\Omega,\mathscr{F})$ 上有界或非负随机变量 Y 与 $x\in E$, $t\in\mathsf{T}$, $t>0$, 有

$$\mathbb{E}^x(Y\circ\theta_t|\mathscr{F}_t)=\mathbb{E}^{X_t}(Y).$$

此外一 Markov 过程称为是正规的, 如果对任何 $x\in E$, $\{x\}\in\mathscr{E}$ 且 $\mathbb{P}^x(X_0=x)=1$. 右连续的, 正规的时齐 Markov 过程简称为右连续简单 Markov 过程.

下面写一个时齐 Markov 过程时, 我们总意味着以上定义的这样一个过程, 虽然有时我们并不把所有的符号写出, 如我们简单地写 $X=(X_t,\mathscr{F}_t,\mathbb{P}^x)$ 或

$X = (X_t, \mathbb{P}^x)$ 等. 右连续简单 Markov 过程是一个重要概念, 在第四章中要详细讨论. 现在, 我们有下面的定理.

定理 2.3.3 (1) 对完备可分度量空间 E 上的正规转移半群 (P_t), 存在以 E 为状态空间的时齐正规 Markov 过程 X 使得对 $t \in \mathsf{T}, x \in E$ 及 $B \in \mathscr{E}$, 有 $P_t(x,B) = \mathbb{P}^x(X_t \in B)$, 这个过程称为是转移半群 $(\mathbb{P}_t)$ 的一个实现;

(2) 设 $(E,\mathscr{E})$ 是可测空间, X 是以 E 为状态空间的时齐 Markov 过程, 则

$$P_t(x,B) := \mathbb{P}^x(X_t \in B),\ x \in E, B \in \mathscr{E}, t \in \mathsf{T}$$

定义了 $(E,\mathscr{E})$ 上的一个转移半群且过程的正规性蕴含着半群的正规性.

证明 只需证明 (2). 首先验证 (P_t) 满足 Chapman-Kolmogorov 方程, 因为 X_s 在 $\mathbb{P}^x$ 之下的分布是 $P_s(x,\cdot)$, 由 Markov 性,

$$\begin{aligned}
P_{t+s}(x,B) &= \mathbb{P}^x(X_{t+s} \in B) \\
&= \mathbb{E}^x(1_{\{X_t \in B\}} \circ \theta_s) = \mathbb{E}^x(\mathbb{P}^{X_s}(X_t \in B)) \\
&= \int_E \mathbb{P}^y(X_t \in B)\mathbb{P}^x(X_s \in \mathrm{d}y) \\
&= \int_E P_s(x,\mathrm{d}y)P_t(y,B).
\end{aligned}$$

另外过程的正规性显然蕴含着半群的正规性. □

例 2.3.2 设 G 是局部紧 Abel 群, 例如 Euclid 空间, G 上的概率测度族 $\pi = \{\pi_t : t > 0\}$ 称为对卷积有半群性, 如果 $\pi_t * \pi_s = \pi_{t+s},\ t,s > 0$; 称为是卷积半群, 如果进一步地当 $t \downarrow 0$ 时, π_t 弱收敛于 ϵ_0, 其中 0 是 G 的单位. 在 Euclid 空间上就有许多重要的卷积半群, 具体的我们将在 §4.3 中论述. 如果 π 是卷积半群, 那么它平凡地诱导一个转移半群 (P_t), 它定义为 $P_t(x,A) := \pi_t(A-x),\ t > 0, x \in G, A \in \mathscr{B}(G)$. 设其对应的正规 Markov 过程为 X, 那么 X 是平稳独立增量过程, 精确地说, X 满足对任何 $t > s > 0$, $A \in \mathscr{B}(G)$, $x \in G$,

$$\mathbb{P}^x(X_t - X_s \in A|\mathscr{F}_s) = \pi_{t-s}(A).$$

注意右边与 x 无关. 事实上,

$$\mathbb{P}^x(X_t - X_0 \in A) = \mathbb{P}^x(X_t \in x + A) = P_t(x, x + A) = \pi_t(A).$$

然后由 Markov 性直接计算得

$$\begin{aligned}\mathbb{P}^x(X_t - X_s \in A|\mathscr{F}_s) &= \mathbb{P}^x((X_{t-s} - X_0)\circ\theta_s \in A|\mathscr{F}_s)\\ &= \mathbb{P}^{X_s}(X_{t-s} - X_0 \in A) = \pi_{t-s}(A).\end{aligned}$$

∎

例 2.3.3 设 E 是有限集合, 一个矩阵 $\mathbf{P} = (p(x,y) : x, y \in E)$ 称为 Markov 矩阵, 如果

(1) 对任何 $x, y \in E$, $p(x,y) \geqslant 0$;

(2) 对任何 $x \in E$, 有 $\sum_{y\in E} p(x,y) = 1$,

那么矩阵 $\mathbf{P}$ 决定 E 上的一个转移半群: 对任何 $x \in E, A \subset E$, 定义核

$$P(x, A) := \sum_{y\in A} p(x,y).$$

那么 P 是 E 上Markov 核, $(P^n : n \geqslant 0)$ 是离散时间转移半群. 所对应的时齐 Markov 过程是有限 Markov 链. 我们将在下节详细讨论. ∎

习　题

1. 设 μ 是 $\mathbf{R}^d$ 上概率测度, 证明: $(x, A) \mapsto \mu(A - x)$ 是一个 Markov 核.
2. 证明: 如果随机过程 X 的转移半群是例 2.3.2 中所构造的 (P_t), 则 X 是一个独立增量过程.
3. 设 $\{\xi_n : n \geqslant 1\}$ 是一个独立随机变量序列, $\{a_n : n \geqslant 1\}$ 是个数列. 定义 $X_0 := 0,\ X_n = \sum_{i=1}^{n} a_i\xi_i$. 证明: $(X_n : n \geqslant 0)$ 是一个具 Markov 性的随机过程.
4. 设 X 是随机过程. 令 $\mathscr{G}_t := \sigma(X_s :\ s \geqslant t)$. 证明: X 的 Markov 性等价于下列两者之一.

   (1) 如果 $A \in \mathscr{G}_t,\ B \in \mathscr{F}_t$, 则

$$\mathbb{P}(A \cap B|X_t) = \mathbb{P}(A|X_t)\mathbb{P}(B|X_t);$$

(2) 如果 $B \in \mathscr{F}_t$, 则 $\mathbb{P}(B|\mathscr{G}_t) = \mathbb{P}(B|X_t)$.

因此, 对于一个 Markov 过程而言, 将来与过去是对称的.

5. 设 $P_t(x,\cdot)$ 是一个均值为 $m_t x$, 方差为 σ_t^2 的正态分布, 问: 当 m_t 与 σ_t^2 满足什么条件时, (P_t) 满足 Chapman-Kolmogorov 方程?

§2.4　Markov链

设 E 是一个有限集或者可列集, 自然地, E 上的拓扑是离散拓扑, σ- 代数就取 E 的全体子集. $E\times E$ 上的函数 $(x,y)\mapsto p_{x,y}$ 称为是 E 上的(保守的)转移函数, 如果它满足:

(1) $p_{x,y} \geqslant 0$;

(2) $\sum\limits_{y\in E} p_{x,y} = 1$.

需要时 $p_{x,y}$ 可写为 $p(x,y)$. 另外我们也将 $p_{x,y}$ 看作一个矩阵 $\mathbf{P}$ 在位置 (x,y) 上的元素, 即

$$\boldsymbol{P} := (p_{x,y} : x, y \in E),$$

那么 $\boldsymbol{P}$ 称为是 E 上的转移矩阵(transition matrix). 这样, 转移函数与转移矩阵是一一对应的. 注意它可能是一个无限行与列的矩阵, 但是它的元素是非负的. 且每行的和是 1.

转移函数 p 完全决定了 E 上的一个 Markov 核

$$P(x,A) := \sum_{y\in A} p_{x,y},$$

$x\in E, A\subset E$. 因此我们说转移函数时也是指它所决定的核. 容易验证, 如果 $\boldsymbol{P}_1, \boldsymbol{P}_2$ 是对应于转移函数 p_1, p_2 的转移矩阵, 对应的核分别为 P_1, P_2, 那么矩阵 $\boldsymbol{P}_1\boldsymbol{P}_2$ 在位置 (x,y) 的元素为 $\sum\limits_{z\in E} p_1(x,z)p_2(z,y)$, 恰好是核的乘积 $P_1P_2(x,\{y\})$, 即 $\boldsymbol{P}_1\boldsymbol{P}_2$ 是对应于 P_1P_2 的转移矩阵, 换句话说核的乘法与矩阵的乘法是一致的. 这样, 我们不区分转移函数, 对应的 Markov 核及其对应的转移矩阵.

设 $\boldsymbol{P} = (p_{x,y} : x, y \in E)$ 是一个转移矩阵, 对任何 $n \geqslant 0$, 定义 $\boldsymbol{P}_n := \boldsymbol{P}^n$, $\boldsymbol{P}_0$ 表示单位矩阵, 用 $p_{x,y}^{(n)}$ 表示矩阵 $\boldsymbol{P}_n$ 位置 (x, y) 的元素, 那么任何 $\boldsymbol{P}_n$ 是转移矩阵, 且 $(\boldsymbol{P}_n : n \geqslant 0)$ 是 E 上的一个转移半群,自然地对任何 $n, m \geqslant 0$, $x, y \in E$,

$$p_{x,y}^{(n+m)} = \sum_{z \in E} p_{x,z}^{(n)} p_{z,y}^{(m)},$$

此方程就是 Chapman-Kolmogorov 方程. 研究转移矩阵的一个重要工具是 Markov 链.

定义 2.4.1 一个 Markov 链是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的一个以 E 为状态空间的随机序列 $\{X_n : n \geqslant 0\}$, 它满足:

(1) Markov 性: 对任何 $n \geqslant 0$, $x_0, x_1, \cdots, x_{n-1}, x, y \in E$, 有

$$\begin{aligned} &\mathbb{P}(X_{n+1} = y | X_n = x, X_{n-1} = x_{n-1}, \cdots, X_1 = x_1, X_0 = x_0) \\ &= \mathbb{P}(X_{n+1} = y | X_n = x); \end{aligned}$$

(2) 时间齐性: 对任何 $n \geqslant 0$, $x, y \in E$, 条件概率 $\mathbb{P}(X_{n+1} = y | X_n = x)$ 与时间 n 无关.

注意关于条件概率的等式只要在两边都有意义时成立就可以了. 但是, (1) 实际上等价于

$$\mathbb{P}(X_{n+1} = y | X_n, \cdots, X_0) = \mathbb{P}(X_{n+1} = y | X_n).$$

容易验证, $p_{x,y} := \mathbb{P}(X_{n+1} = y | X_n = x)$ 是 E 上的转移函数, 再令 $\mu_x := \mathbb{P}(X_0 = x)$. 这时我们说: 在概率 $\mathbb{P}$ 下, X 是一个转移函数是 $p = (p_{x,y})$, 初始分布为 $\mu = (\mu_x)$ 的 Markov 链.

由上一节的定理 2.3.3 立刻推出下列定理, 它说明任何转移函数总对应一个 Markov 链. 称它是保守的, 如果转移函数是保守的.

定理 2.4.1 $\boldsymbol{P} = (p_{x,y} : x, y \in E)$ 是 E 上的一个转移矩阵当且仅当存在一个以 E 为状态空间, 以 $\{0, 1, 2, \cdots\}$ 为时间参数集的 Markov 过程

$$X = (\Omega, \mathscr{F}, (\mathbb{P}^x : x \in E), (X_n : n \geqslant 0), (\theta_n : n \geqslant 0)),$$

使得对任何 $n \geqslant 0$, $p^{(n)}(x, y) = \mathbb{P}^x(X_n = y)$.

定理说明在概率 $\mathbb{P}^x$ 下, X 是转移函数为 p, 初始分布为 x 的单点测度(或说从 x 出发)的 Markov 链. 精确地说, $(\mathbb{P}^x : x \in E)$ 是 $(\Omega, \mathscr{F})$ 上的概率, $(X_n : n \geqslant 0)$ 是随机序列, θ_k 是 Ω 上推移算子, 即满足 $X_n \circ \theta_k = X_{n+k}$, $n, k \geqslant 0$, (实际上, 这里我们只需要有一个 θ 满足 $X_n \circ \theta = X_{n+1}$, $n \geqslant 0$ 就够了.) 它们满足:

(1) $\mathbb{P}^x(X_0 = x) = 1$;

(2) Markov 性: 对任何 $n \geqslant 0$, $B \in \mathscr{F}_\infty$, $x \in E$, 有

$$\mathbb{E}^x(1_B \circ \theta_n | \mathscr{F}_n) = \mathbb{E}^{X_n}(B),$$

其中 $\mathscr{F}_n := \sigma(\{X_i : 0 \leqslant i \leq n\})$, 而 $\mathscr{F}_\infty := \sigma(X_i : i \geqslant 0)$.

注 有两个问题需要说明. 第一, Markov 性也可以如定义 2.4.1 那样用条件概率的方式描述. 第二, 上面的时间 n 可以用一类随机时间代替, 也就是强 Markov 性成立. 一个映射 $\tau : \Omega \longrightarrow \mathbf{Z}^+ \cup \{+\infty\}$ 称为是一个停时, 如果对任何 $n \geqslant 0$, 有 $\{\tau \leqslant n\} \in \mathscr{F}_n$, 这等价于对任何 $n \geqslant 0$, $\{\tau = n\} \in \mathscr{F}_n$. 注意, τ 可以取 $+\infty$ 为值, 且一定是 $\mathscr{F}_\infty$ (广义)随机变量. 容易验证固定时间 n 是一个停时. 定义 $\mathscr{F}_\tau$ 是 $\mathscr{F}_\infty$ 中满足对任何 $n \geqslant 0$, $A \cap \{\tau \leqslant n\} \in \mathscr{F}_n$ 的元素 A 的全体, 那么

(1) $\mathscr{F}_\tau$ 是一个 σ- 代数;

(2) 当 $\tau \equiv n$ 时, $\mathscr{F}_\tau = \mathscr{F}_n$. (请读者自行验证.) 再定义停止位置 $X_\tau := X_n$ 当 $\tau = n$ 时, $n \geqslant 0$. 这样 X_τ 在集合 $\{\tau < \infty\}$ 上被定义, 且在此集合上是 $\mathscr{F}_\tau$ 可测的. 类似定义 $\theta_\tau := \theta_n$ 当 $\tau = n$ 时. 同样 θ_τ 也在集合 $\{\tau < \infty\}$ 上被定义. X 有下面的强 Markov 性: 设 τ 是一个停时, 则对任何 $B \in \mathscr{F}_\infty$, $x \in E$, 有

$$\mathbb{E}^x(1_B \circ \theta_\tau; \tau < +\infty | \mathscr{F}_\tau) = \mathbb{E}^{X_\tau}(1_B) 1_{\{\tau < +\infty\}}.$$

证明是简单的. 任取 $A \in \mathscr{F}_\tau$, 那么对任何 $n \geqslant 0$, $A \cap \{\tau = n\} \in \mathscr{F}_n$, 由 Markov 性,

$$\begin{aligned}
\mathbb{E}^x(1_B \circ \theta_\tau; A \cap \{\tau < \infty\}) &= \sum_{n=0}^{\infty} \mathbb{E}^x(1_B \circ \theta_\tau; A \cap \{\tau = n\}) \\
&= \sum_{n=0}^{\infty} \mathbb{E}^x(1_B \circ \theta_n; A \cap \{\tau = n\}) \\
&= \sum_{n=0}^{\infty} \mathbb{E}^x(\mathbb{E}^{X_n}(1_B); A \cap \{\tau = n\})
\end{aligned}$$

$$= \mathbb{E}^x(\mathbb{E}^{X_\tau}(1_B); A \cap \{\tau < \infty\}).$$

这就证明了强 Markov 性. 因此在 Markov 链的情形下, Markov 性与强 Markov 性是等价的, 但强 Markov 性是一个很有用的工具.

例 2.4.1 (Bernoulli-Laplace 扩散模型) 设 A, B 两箱中各有 r 个球, 其中 r 个白, r 个黑. 记 X_0 是开始时 A 箱中白球个数, 然后各任取一球对换, X_n 是经 n 次对换后 A 箱中白球个数. $X = (X_n : n \geqslant 1)$ 是随机序列, 状态空间是 $E = \{0, 1, 2, \cdots, n\}$. 容易验证 X 是 Markov 链且

$$p_{x,x-1} = \left(\frac{x}{r}\right)^2, \ p_{x,x} = 2\frac{i(r-x)}{r^2}, \ p_{x,x+1} = \left(\frac{r-x}{r}\right)^2.$$

这是一个关于两种液体混合的概率模型. ∎

例 2.4.2 (无限制随机游动) 例 2.1.1 中的随机序列 $S = (S_n : n \geqslant 0)$ 是一个 Markov 链. 转移矩阵为

$$\boldsymbol{P} := \begin{bmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & q & 0 & p & 0 & 0 & \cdots \\ \cdots & 0 & q & 0 & p & 0 & \cdots \\ \cdots & 0 & 0 & q & 0 & p & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}.$$

∎

例 2.4.3 (具吸收壁的随机游动) 设 Markov 链 X 有状态空间 $E = \{0, 1, \cdots, r\}$, 转移矩阵

$$\boldsymbol{P} := \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & q & 0 & p & \cdots & 0 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & q & 0 & p & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \end{bmatrix},$$

$p+q=1$. 直观地, 当 X_n 位于状态 1 与 $r-1$ 之间时, 则下一步向左的概率为 q, 向右的概率为 p, 而当 X_n 达到边界 0 或 r 时, 它将不再离开.

例 2.4.4　A, B 两人比赛, A 赢的概率是 p, 输的概率是 $q=1-p$.

(1) 如果规则是比赛到某人胜出两场为赢, 求 A 赢的概率;

(2) 如果规则是某人连续赢两场为赢,求 A 赢的概率.

对规则 (1). 用 X_n 表示到第 n 局结束时, A 的净赢局数. 显然 X_n 是个具吸收壁 Markov 链, 状态依次为 $-2, -1, 0, 1, 2$, 其转移矩阵为

$$\boldsymbol{P}:=\begin{bmatrix}1&0&0&0&0\\q&0&p&0&0\\0&q&0&p&0\\0&0&q&0&p\\0&0&0&0&1\end{bmatrix}.$$

用 $p^A(x)$ 表示从状态 x 开始最终 A 赢的概率. 那么$p^A(-2)=0$, $p^A(2)=1$. 由全概率公式

$$\begin{aligned}p^A(-1)&=p\cdot p^A(0),\\p^A(0)&=q\cdot p^A(-1)+p\cdot p^A(1),\\p^A(1)&=q\cdot p^A(0)+p,\end{aligned}$$

可以容易地推出 A 最终赢的概率是 $p^A(0)=\dfrac{p^2}{1-2pq}$.

对规则 (2). 这个 Markov 链不能解决问题. 让我们用 $X_n=A$ (或 $X_n=B$)表示第 n 局 A 赢(B 赢). $Y_n=(X_{n-1}, X_n)$, $n\geqslant 2$, 则 Y_n 有4个状态依次为 AA, AB, BA, BB. 转移矩阵为

$$\boldsymbol{P}:=\begin{bmatrix}1&0&0&0\\0&0&p&q\\p&q&0&0\\0&0&0&1\end{bmatrix}.$$

显然 $p^A(AA)=1$, $p^A(BB)=0$, $p^A(AB)=0+p\cdot p^A(BA)$, $p^A(BA)=p+q\cdot p^A(AB)$. 先观察头两局比赛, 由全概率公式,

$$p_2^A=p^2+pq\cdot p^A(AB)+qp\cdot p^A(BA)=\frac{p^2(1+q)}{1-pq}.$$

例 2.4.5 (对称随机游动) 设 E 是 $\mathbf{R}^d$ 上的整数格点, 任意 $x\in E$ 都有 $2d$ 个邻居(与 x 的距离等于 1), 如果 y 是 x 的邻居, 定义 $p_{x,y}:=\dfrac{1}{2d}$, 否则定义 $p_{x,y}:=0$, 则 $\boldsymbol{P}=(p_{x,y}:x,y\in E)$ 是一个转移矩阵. 一个具有此转移矩阵的 Markov 链称为是对称随机游动.

例 2.4.6 (与图上简单随机游动) 设 E 是一个简单图的顶点集合, 对任何 $x,y\in E$, 若 x,y 有线直接连接, 则令 $S(x,y)=1$ 否则 $S(x,y)=0$, 则 S 是对称的. 设

$$p_{x,y}=\frac{S(x,y)}{\sum\limits_{z\in E}S(x,z)},$$

则 $(p_{x,y}:x,y\in E)$ 定义了 E 上一个转移函数, 对应的 Markov 链称为是图上简单随机游动. 上例中的对称随机游动就是整数格点图上简单随机游动.

现在设 $(p_{x,y}:x,y\in E)$ 是转移函数, X 是对应的 Markov 链. 设 $x,y\in E$, 称 x 可达 y, 如果存在 $n\geqslant 0$, 使 $p_{x,y}^{(n)}>0$, 称 x,y 互达, 如果 x 可达 y 且 y 可达 x. 由 Chapman-Kolmogorov 方程, 状态间互达关系是等价关系. 称 X 不可分, 如果 E 的任何两个状态可互达. 对 $y\in E$, 令 $\tau_y:=\inf\{n\geqslant 1:X_n=y\}$ (约定 $\inf\varnothing=\infty$), 容易验证 τ_y 是一个停时, 它被称为是状态 y 的首达时. 这是一个非常基本的停时, 有许多性质和应用. 对 $n\geqslant 0$, 令 $f_{x,y}^{(n)}:=\mathbb{P}^x(\tau_y=n)$, 显然 $f_{x,y}^{(n)}\leqslant p_{x,y}^{(n)}$. 因为 $\{\tau_y=k\}\in\mathscr{F}_k$, $X_{\tau_y}=y$, 故对 $n\geqslant 1$,

$$\begin{aligned}p_{x,y}^{(n)}&=\mathbb{P}^x(\{X_n=y\})=\mathbb{P}^x(X_n=y,\tau_y\leqslant n)\\&=\sum_{k=1}^n\mathbb{P}^x(X_{n-k}\circ\theta_k=y,\tau_y=k)\\&=\sum_{k=1}^n\mathbb{E}^x(\mathbb{P}^{X_k}(X_{n-k}=y);\tau_y=k)\end{aligned}$$

$$= \sum_{k=1}^{n} \mathbb{P}^x(\tau_y = k)\mathbb{P}^y(X_{n-k} = y)$$

$$= \sum_{k=1}^{n} f_{x,y}^{(k)} p_{y,y}^{(n-k)}.$$

用 $P_{x,y}(t)$, $F_{x,y}(t)$ 分别表示 $(p_{x,y}^{(n)} : n \geqslant 0)$, $(f_{x,y}^{(n)} : n \geqslant 0)$ 的母函数. 将上面关于转移概率与首达概率的重要公式用母函数的语言写出来为

$$P_{x,y}(t) = 1_{\{x=y\}} + F_{x,y}(t)P_{y,y}(t).$$

再定义

$$f_{x,y} := \mathbb{P}^x(\tau_y < \infty) = \sum_{n \geqslant 1} f_{x,y}^{(n)},$$

这是过程从状态 x 经有限步到达 y 的概率, 可以看出当 $x \neq y$ 时, x 可达 y 当且仅当 $f_{x,y} > 0$.

一个状态 $x \in E$ 称为是常返的, 如果从 x 出发以概率 1 回到 x, 即 $f_{x,x} = 1$, 否则 x 称为暂留的. Markov 链 X 称为是常返的, 如果其所有状态常返; 称为暂留如果其所有状态暂留.

定理 2.4.2　我们有下列准则:

(1) 状态 x 是常返的等价于 $\mathbb{P}^x(\limsup\{X_n = x\}) = 1$ 也等价于 $\sum\limits_n p_{x,x}^{(n)} = \infty$;

(2) 状态 x 是暂留的等价于 $\mathbb{P}^x(\limsup\{X_n = x\}) = 0$ 也等价于 $\sum\limits_n p_{x,x}^{(n)} < \infty$.

证明　对 $y \in E$, 令

$$\tau^{(0)} := 0,\ \tau := \tau_y,$$

$$\tau^{(k)} := \tau^{(k-1)} + \tau\circ\theta_{\tau^{(k-1)}},\ k \geqslant 1,$$

那么 $\tau^{(k)}$ 实际上是 X 第 k 次遇到 y 的时刻, 它也是一个停时. 再定义 $A_k(y) := \{\tau^{(k)} < \infty\}$, 即表示事件: X 至少有 k 次到达 y, 则 $\{A_k(y) : k \geqslant 1\}$ 是单降集列, 且

$$\limsup_n\{X_n = y\} = \bigcap_k A_k(y),$$

由强 Markov 性且因为 $X_{\tau^{(k-1)}} = y$,

$$\begin{aligned}\mathbb{P}^x(A_k(y)) &= \mathbb{P}^x(\tau^{(k)} < \infty) = \mathbb{P}^x(\tau^{(k-1)} < \infty, \tau\circ\theta_{\tau^{(k-1)}} < \infty)\\ &= \mathbb{E}^x(\mathbb{P}^{X(\tau^{(k-1)})}(\tau < \infty); \tau^{(k-1)} < \infty)\\ &= f_{y,y}\mathbb{P}^x(\tau^{(k-1)} < \infty)\\ &= \cdots = f_{x,y}(f_{y,y})^{k-1}.\end{aligned}$$

因此, 当 $y = x$ 时,

$$\mathbb{P}^x(\limsup_n\{X_n = x\}) = \lim_k (f_{x,x})^k = \begin{cases}0, & f_{x,x} < 1,\\ 1, & f_{x,x} = 1.\end{cases}$$

又由 Borel-Cantelli 引理, $\sum\limits_n p_{x,x}^{(n)} < \infty$ 蕴含着

$$\mathbb{P}^x(\limsup\{X_n = x\}) = 0,$$

最后, 由转移概率与首达概率的公式得 $\sum\limits_n p_{x,x}^{(n)} = 1 + f_{x,x}\sum\limits_n p_{x,x}^{(n)}$, 因此, $f_{x,x} < 1$ 当且仅当 $\sum\limits_n p_{x,x}^{(n)} < \infty$. □

下列定理表明在一个互达等价类中, 或者所有状态是常返的, 或者所有状态是暂留的.

定理 2.4.3 如果 $x, y \in E$ 是互达的, 则下列两者之一成立:

(1) x, y 都是常返的, $\mathbb{P}^x(\limsup\{X_n = y\}) = 1$ 且 $\sum\limits_n p_{x,y}^{(n)} = \infty$;

(2) x, y 都是暂留的, $\mathbb{P}^x(\limsup\{X_n = y\}) = 0$ 且 $\sum\limits_n p_{x,y}^{(n)} < \infty$.

证明 设 $x \neq y$. 因 x, y 互达, 存在 $i, j \geqslant 1$, 使得 $p_{x,y}^{(i)} \cdot p_{y,x}^{(j)} > 0$, 而对任何 $n \geqslant 1$,

$$p_{x,x}^{(i+n+j)} \geqslant p_{x,y}^{(i)}p_{y,y}^{(n)}p_{y,x}^{(j)},\ p_{y,y}^{(i+n+j)} \geqslant p_{y,x}^{(j)}p_{x,x}^{(n)}p_{x,y}^{(i)},$$

因此定理 2.4.2 推出 x, y 或都是常返或都是暂留. 由转移与首达概率公式, $\sum\limits_{n\geqslant 0} p_{x,y}^{(n)} = f_{x,y}\sum\limits_{n\geqslant 0} p_{y,y}^{(n)}$. 因 $f_{x,y} > 0$, 故 $\sum\limits_n p_{x,y}^{(n)} < \infty$ 当且仅当 $\sum\limits_n p_{y,y}^{(n)} < \infty$. 由强 Markov 性,

$$\begin{aligned}\mathbb{P}^x(\limsup\{X_n=y\}) &= \mathbb{P}^x(\limsup\{X_n=y\},\tau_y<\infty)\\ &= \mathbb{E}^x(1_{\limsup\{X_n=y\}}\circ\theta_{\tau_y};\tau_y<\infty)\\ &= \mathbb{E}^x(\mathbb{P}^y(\limsup\{X_n=y\});\tau_y<\infty)\\ &= f_{x,y}\mathbb{P}^y(\limsup\{X_n=y\}),\end{aligned}$$

因此若 x,y 是暂留的, 则 $\mathbb{P}^x(\limsup\{X_n=y\})=0$; 若 x,y 是常返的, 则

$$\mathbb{P}^x(\limsup\{X_n=y\})=f_{x,y},$$

故我们仅需验证 $f_{x,y}=1$ 即可.

事实上, 设 y 常返, 对任何的 $k\geqslant 1$, $\mathbb{P}^y(\limsup\{X_n=y\})=1$ 蕴含着 $\mathbb{P}^y(\bigcup_{n>k}\{X_n=y\})=1$ 等价于 $\mathbb{P}^y(k<\tau_y<\infty)=1$, 容易验证 $\tau_y\circ\theta_k+k=\inf\{n>k:X_n=y\}$, 故 $\{k<\tau_y<\infty\}=\{\tau_y\circ\theta_k<\infty\}$. 由 Markov 性得

$$\begin{aligned}p_{y,x}^{(k)} &= \mathbb{P}^y(X_k=x)\\ &= \mathbb{P}^y(X_k=x,\tau_y\circ\theta_k<\infty)\\ &= \mathbb{E}^y(\mathbb{P}^y(X_k=x,\tau_y\circ\theta_k<\infty|\mathscr{F}_k))\\ &= \mathbb{E}^y(\mathbb{P}^{X_k}(\tau_y<\infty);X_k=x)\\ &= f_{x,y}p_{y,x}^{(k)}.\end{aligned}$$

由 y 可达 x 推出 $f_{x,y}=1$. □

子集 $C\subset E$ 称为是闭的, 如果对任何 $x\in C$, $y\notin C$ 有 $p_{x,y}=0$. 等价于对任何 $x\in C$, 有 $\sum\limits_{y\in C}p_{x,y}=1$. 闭集中的任何状态不能到达闭集外的状态. 那么 X 不可分当且仅当 E 没有非平凡闭子集(习题). 上面证明的最后实际上证明了, 如果 x 可达 y 且 x 常返, 那么 y 也可达 x 且 y 也常返. 因此所有 X 的常返状态全体是个闭子集.

例 2.4.7 让我们考虑例 2.4.2 中的无限制随机游动. 显然对所有 x, 从 x 出发经 $2n$ 步回到 x 的概率 $p_{x,x}^{(2n)}=\binom{2n}{n}p^nq^n$, 当然 $p_{x,x}^{(2n+1)}=0$. 由 Stirling 公式:

$n! \sim \dfrac{n^n}{e^n}\sqrt{2\pi n}$, 得 $\dbinom{2n}{n} \sim \dfrac{4^n}{\sqrt{\pi n}}$, 故

$$p_{x,x}^{(n)} \sim \frac{4^n}{\sqrt{\pi n}}(pq)^n.$$

因此当 $p = q = \dfrac{1}{2}$ 时, $\sum_n p_{x,x}^{(n)} = \infty$, 所有状态常返, 即 X 常返; 否则所有状态暂留, 即 X 暂留. 实际上, 我们可以计算 $f_{0,0}$. 由展开式

$$\frac{1}{\sqrt{1-4t}} = \sum_{n=0}^{\infty}\binom{2n}{n}t^n$$

推出 $\{p_{0,0}^{(n)}\}$ 的母函数为

$$P(t) = \frac{1}{\sqrt{1-4pqt^2}}.$$

因此 $\{f_{0,0}^{(n)}\}$ 的母函数为

$$F(t) = 1 - \sqrt{1-4pqt^2}.$$

推出 $\mathbb{P}^0(\tau_0 < \infty) = f_{0,0} = F(1) = 1 - |p-q|$. ▮

定理 2.4.4 (Polya)　$\mathbf{R}^d$ 上的对称随机游动当 $d = 1, 2$ 时是常返的, 而当 $d \geqslant 3$ 时是暂留的.

证明　显然 n 返回的概率与维数 d 有关而与状态无关, 记为 $q_n^{(d)}$. $d = 1$ 的情况已在例 2.4.7 中证明. 当 $d = 2$ 时, 同样奇数步返回的概率是零, 偶数 $2n$ 返回必定是 k 步向上, k 步向下, $n-k$ 步向左, $n-k$ 步向右, 因此

$$\begin{aligned} q_{2n}^{(2)} &= \frac{1}{4^{2n}}\sum_{k=0}^{n}\frac{(2n)!}{k!k!(n-k)!(n-k)!} \\ &= \frac{1}{4^{2n}}\binom{2n}{n}\sum_{k=0}^{n}\binom{n}{k}\binom{n}{n-k} \\ &= \frac{1}{4^{2n}}\binom{2n}{n}^2 \sim \frac{1}{\pi n}. \end{aligned}$$

因此 $\sum\limits_n q_n^{(2)} = \infty$.

当 $d=3$ 时, 由同样的分析得

$$\begin{aligned} q_{2n}^{(3)} &= \frac{1}{6^{2n}} \sum_{k_1+k_2+k_3=n} \frac{(2n)!}{k_1!k_1!k_2!k_2!k_3!k_3!} \\ &= \frac{1}{6^{2n}} \binom{2n}{n} \sum_{k_1+k_2+k_3=n} \left[\frac{n!}{k_1!k_2!k_3!}\right]^2 \\ &= \frac{1}{2^{2n}} \binom{2n}{n} \sum_{k_1+k_2+k_3=n} \left[\frac{1}{3^n}\frac{n!}{k_1!k_2!k_3!}\right]^2. \end{aligned}$$

中括号中是一个三项分布, 其和为 1. 因此

$$\begin{aligned} &\sum_{k_1+k_2+k_3=n} \left[\frac{1}{3^n}\frac{n!}{k_1!k_2!k_3!}\right]^2 \\ &\qquad \leqslant \frac{1}{3^n} \cdot \max\left\{\frac{n!}{k_1!k_2!k_3!}:\ k_1+k_2+k_3=n\right\}. \end{aligned}$$

右边最大在 $k_1=k_2=k_3$ 或最接近处达到, 由 Stirling 公式, 最大值与 $3^n \cdot n^{-1}$ 同级. 因此 $q_{2n}^{(3)}$ 与 $n^{-\frac{3}{2}}$ 同级, 故 $\sum\limits_n q_n^{(3)} < \infty$, 即 3 维对称随机游动是暂留的. 当 $d \geqslant 4$ 时, 证明类似. □

对 $x \in E$, 定义 x 的周期, 记 $d(x)$, 为集合 $\{n \geqslant 1 : p_{x,x}^{(n)} > 0\}$ 的最大公因数, (如集合是空, $d(x) := 0$.) 若 $d(x)=1$, 称 x 是非周期的. 称 Markov 链是非周期的, 如果其所有状态都是非周期的.

引理 2.4.1 (1) 如果 x, y 互达, 则 $d(x)=d(y)$;

(2) 如果 X 是不可分非周期的 Markov 链, 则对 $x, y \in E$,存在 n_0, 使当 $n > n_0$ 时, $p_{x,y}^{(n)} > 0$.

证明 (1) 存在 $s, t \in \mathsf{T}$, 使得 $p_{x,y}^{(s)} p_{y,x}^{(t)} > 0$, 对 $n \geqslant 1$, 如果 $p_{x,x}^{(n)} > 0$, 则 $p_{y,y}^{(t+n+s)} \geqslant p_{y,x}^{(t)} p_{x,x}^{(n)} p_{x,y}^{(s)} > 0$, 故 $d(y)$ 整除 $s+n+t$, 而这时 $p_{x,x}^{(n)} > 0$, 故 $d(y)$ 也整除 $s+2n+t$, 因此 $d(y)$ 整除 n, 即 $d(y) \leqslant d(x)$. 同理有 $d(x) \leqslant d(y)$.

(2) 设 $G_y := \{n \geqslant 1 : p_{y,y}^{(n)} > 0\}$, 则 G_y 对于加法封闭, 且因 G_y 的最大公因数是 1, 故存在 m_0, 使 G_y 包含比 m_0 大的所有自然数. (这是一个数论问题, 读者可先利用 G_y 的最大公因子是 1 的事实证明: 存在 n, 使得 $n, n+1 \in G_y$, 然后利用对加法封闭的性质证明结论.) 另外, 存在 $s \in \mathsf{T}$, 使 $p_{x,y}^{(s)} > 0$, 故对任何 $n > n_0 := m_0 + s$, $p_{x,y}^{(n)} > 0$. □

E 上的一个概率分布 $(\pi_y : y \in E)$ 称为是转移矩阵 $\boldsymbol{P}$ 的平稳分布, 如果

$$\sum_{x \in E} \pi_x p_{x,y} = \pi_y, \ y \in E.$$

定理 2.4.5 如果一个不可分非周期的 Markov 链 X 存在有平稳分布 $(\pi_x : x \in E)$, 那么

(1) X 是常返的;

(2) 对任何 $x, y \in E$, $\lim_n p_{x,y}^{(n)} = \pi_y$, 平稳分布是唯一的;

(3) 对任何 $x \in E$, $\pi_x > 0$.

证明 (1) 由平稳分布的定义, 对任何 $n \geqslant 1$,

$$\sum_{x \in E} \pi_x p_{x,y}^{(n)} = \pi_y, \ y \in E.$$

假设 X 是暂留的, 则 $\lim_n p_{x,y}^{(n)} = 0$, 由控制收敛定理, 对任何 $y \in E$, $\pi_y = 0$, 与 (π_y) 是概率分布矛盾.

(2) 我们首先构造 $E \times E$ 上的转移矩阵, 对 $(x,y),(u,v) \in E \times E$, 定义

$$p((x,y),(u,v)) := p_{x,u} p_{y,v},$$

容易验证 $(p((x,y),(u,v)) : (x,y),(u,v) \in E \times E)$ 是 $E \times E$ 上的转移矩阵, 称为是 $\mathbb{P}$ 的耦合矩阵, 存在以 $E \times E$ 为状态空间的 Markov 链 $((Y_n, Z_n) : n \in \mathsf{T})$ (对某个给定的初始分布), 以耦合矩阵为转移矩阵, 称其为耦合链, 显然 (Y_n) 与 (Z_n) 是两个独立的 Markov 链, 因此

$$p^{(n)}((x,y),(u,v)) = p_{x,u}^{(n)} p_{y,v}^{(n)},$$

由引理 2.4.1, 耦合链也是不可分非周期的, 且 $\pi_{(x,y)} := \pi_x\pi_y,\ x,y \in E\times E$ 是耦合链的平稳分布, 故耦合链也是常返的. 固定 $z\in E$, 令

$$\tau := \inf\{n \geqslant 1:\ (Y_n, Z_n) = (z,z)\},$$

即首次命中 (z,z) 的时刻, 则对 $m \leqslant n$,

$$\begin{aligned}
&\mathbb{P}^{(x,y)}(\{(Y_n,Z_n)=(u,v),\tau=m\})\\
&\quad=\mathbb{P}^{(x,y)}(\{(Y_n,Z_n)=(u,v)|\tau=m)\mathbb{P}^{(x,y)}(\tau=m)\\
&\quad=\mathbb{P}^{(z,z)}(\{(Y_{n-m},Z_{n-m})=(u,v)\})\mathbb{P}^{(x,y)}(\tau=m)\\
&\quad=p_{z,u}^{(n-m)}p_{z,v}^{(n-m)}\mathbb{P}^{(x,y)}(\tau=m).
\end{aligned}$$

由此推出

$$\mathbb{P}^{(x,y)}(Y_n=u,\tau=m)=\mathbb{P}^{(x,y)}(Z_n=u,\tau=m).$$

因此

$$\begin{aligned}
p_{x,u}^{(n)}=\mathbb{P}^{(x,y)}(Y_n=u)&\leqslant\mathbb{P}^{(x,y)}(Y_n=u,\tau\leqslant n)+\mathbb{P}^{(x,y)}(\tau>n)\\
&=\mathbb{P}^{(x,y)}(Z_n=u,\tau\le n)+\mathbb{P}^{(x,y)}(\tau>n)\\
&\leqslant\mathbb{P}^{(x,y)}(Z_n=u)+\mathbb{P}^{(x,y)}(\tau>n)\\
&=p_{y,u}^{(n)}+\mathbb{P}^{(x,y)}(\tau>n).
\end{aligned}$$

同理可证

$$p_{y,u}^{(n)}\leqslant p_{x,u}^{(n)}+\mathbb{P}^{(x,y)}(\tau>n).$$

故 $|p_{x,u}^{(n)}-p_{y,u}^{(n)}|\leqslant\mathbb{P}^{(x,y)}(\tau>n)$, 而耦合链是常返的蕴含着 $\mathbb{P}^{(x,y)}(\tau<\infty)=1$, 因此推出

$$\lim|p_{x,z}^{(n)}-p_{y,z}^{(n)}|=0,\ x,y,z\in E.$$

另一方面

$$\pi_z-p_{x,z}^{(n)}=\sum_{y\in E}\pi_y p_{y,z}^{(n)}-p_{x,z}^{(n)}=\sum_{y\in E}\pi_y(p_{y,z}^{(n)}-p_{x,z}^{(n)}),$$

由控制收敛定理, $\lim_n p_{x,z}^{(n)}=\pi_z,\ x,z\in E$.

(3) 因 x, y 互达, 存在 $i, j \geqslant 1$ 使得 $p_{x,y}^{(i)} \cdot p_{y,x}^{(j)} > 0$, 而对任何 $n \in \mathsf{T}$,

$$p_{x,x}^{(i+n+j)} \geqslant p_{x,y}^{(i)} p_{y,y}^{(n)} p_{y,x}^{(j)},$$

令 $n \to \infty$, $\pi_x \geqslant p_{x,y}^{(i)} p_{y,x}^{(j)} \pi_y$, 因此如果某个 π_x 是零, 则所有其他都是零, 故 $\{\pi_x : x \in E\}$ 一定全是严格正的. □

非周期的条件是本质的, 例如我们取 $E = \{0, 1\}$, $p_{0,1} = 1, p_{1,0} = 1$, 则容易验证它是不可分常返链, 平稳分布 $\pi_0 = \pi_1 = \dfrac{1}{2}$, 但 $p_{x,y}^{(n)}$ 极限不存在.

定理 2.4.6 如果一个不可分非周期的 Markov 链 X 没有平稳分布, 则对任何 $x, y \in E$, $\lim_n p_{x,y}^{(n)} = 0$.

证明 我们仅需考虑 X 是常返的情况, 这时如果在上面定理证明中构造的耦合链是暂留的, 则由定理 2.4.3,

$$\sum_n (p_{x,y}^{(n)})^2 = \sum_n p^{(n)}((x,x),(y,y)) < \infty,$$

故 $\lim_n p_{x,y}^{(n)} = 0$.

现在设耦合链是常返的, 假设结论不对, 因 E 是至多可列的, 故存在一个子列 (n_u), 使 $\lim_u p_{x,y}^{(n_u)}$ 对所有 $x, y \in E$ 存在, 且不全为零, 由上面定理的证明中可知极限值与 x 无关, 令 $\alpha_y := \lim_u p_{x,y}^{(n_u)}$. 对 E 的任意有限子集 M,

$$\sum_{y \in M} \alpha_y = \lim_u \sum_{y \in M} p_{x,y}^{(n_u)} \leqslant 1,$$

故 $\alpha := \sum\limits_{y \in E} \alpha_y \leqslant 1$. 另外,

$$\sum_{z \in M} p_{x,z}^{(n_u)} p_{z,y} \leqslant p_{x,y}^{(n_u+1)} = \sum_{z \in E} p_{x,z} p_{z,y}^{(n_u)},$$

由控制收敛定理,

$$\sum_{z \in M} \alpha_z p_{z,y} \leqslant \sum_{z \in E} p_{x,z} \alpha_y = \alpha_y,$$

这推出 $\sum\limits_{z \in E} \alpha_z p_{z,y} \leqslant \alpha_y$. 如果其中有一个严格小于, 则

$$\alpha = \sum_{y\in E}\alpha_y > \sum_{y\in E}\sum_{z\in E}\alpha_z p_{z,y} = \sum_{z\in E}\alpha_z = \alpha,$$

导致矛盾, 故对任何 $y \in E$, $\sum_{z\in E}\alpha_z p_{z,y} = \alpha_y$. 因 $\alpha > 0$, 记 $\pi_y := \dfrac{\alpha_y}{\alpha}$, 那么 $(\pi_y : y \in E)$ 是 $(p_{x,y} : x, y \in E)$ 的一个平稳分布, 与条件矛盾. □

为了进一步区别常返的类型, 我们引入正常返与零常返的概念. 设状态 y 是常返的, 称它是正常返的, 如果平均回转时间有限

$$\mu_y := \mathbb{E}^y(\tau_y) = \sum_{n\geqslant 1} n f_{y,y}^{(n)} < \infty,$$

否则称 y 是零常返的.

定理 2.4.7　设 y 是常返的且 $\lim_n p_{y,y}^{(n)}$ 存在记为 u_y, 则 $u_y > 0$ 当且仅当 $\mu_y < \infty$, 这时 $u_y = \dfrac{1}{\mu_y}$.

证明　数列 $(p_{y,y}^{(n)})$ 与 $(f_{y,y}^{(n)})$ 的母函数 $P(t)$ 与 $F(t)$ 满足 $P(t)(1 - F(t)) = 1$. 因为 $\lim_n p_{y,y}^{(n)}$ 存在, 由 Abel 定理(见§1.5习题)或其推论推出

$$\lim_n p_{y,y}^{(n)} = \lim_{t\uparrow 1}\frac{1-t}{1-F(t)} = \frac{1}{F'(1)}.$$

因此定理的结论是显然的. □

从以上结果可以看出, 一个不可分非周期 Markov 链可分为 3 种情况:

(1) 链是暂留的, 这时 $\sum_n p_{y,y}^{(n)} < \infty$, $y \in E$, 不存在平稳分布;

(2) 链是零常返的, 这时 $\sum_n p_{y,y}^{(n)} = \infty$, 但 $\lim_n p_{y,y}^{(n)} = 0$, 不存在平稳分布;

(3) 链是正常返的, 这时 $\sum_n p_{y,y}^{(n)} = \infty$, 且 $\lim_n p_{y,y}^{(n)} = \pi_y > 0$, 存在平稳分布.

例 2.4.8　设 E 是非负整数集, 令

$$\boldsymbol{P} := \begin{bmatrix} q_0 & p_0 & 0 & 0 & \cdots \\ q_1 & 0 & p_1 & 0 & \cdots \\ q_2 & 0 & 0 & p_2 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

其中 p_x, q_x 是正的, 且 $p_x + q_x = 1$, 因此 $\boldsymbol{P}$ 是一个转移矩阵. 设 X 是一个以 $\mathbb{P}$ 为转移矩阵的 Markov 链, 容易验证 X 是不可分非周期的. 对 $n \geqslant 1$, $\sum\limits_{k=1}^{n} f_{0,0}^{(k)} = 1 - \mathbb{P}_0(\tau_0 > n) = 1 - p_0p_1\cdots p_{n-1}$, 当 $n \to \infty$ 时, 乘积 $p_0p_1\cdots p_{n-1}$ 极限存在, 记为 a, 显然 $a \in [0,1]$, 且 X 是常返的当且仅当 $a = 0$.

解方程 $\sum\limits_{x \in E} \pi_x p_{x,y} = \pi_y$, $y \in E$ 得 $\pi_x = \pi_0 p_0 p_1 \cdots p_{x-1}$, $x \geqslant 1$. 因此, 上述 Markov 链有平稳分布当且仅当级数

$$\sum_{x \geqslant 1} p_0 p_1 \cdots p_{x-1}$$

收敛. ▮

习 题

1. 证明: x 可达 y 当且仅当存在 $n \geqslant 1$, $x_i \in E$, $0 \leqslant i \leqslant n$, 使得 $p_{x_{i-1},x_i} > 0$, $1 \leqslant i \leqslant n$ 且 $x_0 = x$, $x_n = y$.
2. 证明: $y \in E$ 是常返的当且仅当对任何 $k \geqslant 1$, 有 $\mathbb{P}^y(\tau_y \circ \theta_k < \infty) = 1$.
3. 证明: 例 2.4.3 中具吸收壁的随机游动必定在有限步达到状态 0 或 r.
4. 证明: 对任何 $x, y \in E$,
$$f_{x,y} \sum_{k=0}^{\infty} p_{y,y}^{(k)} = \sum_{n=1}^{\infty} \sum_{m=1}^{n} f_{x,y}^{(m)} p_{y,y}^{(n-m)} = \sum_{n=1}^{\infty} p_{x,y}^{(n)}.$$
5. 证明: 如果 $y \in E$ 暂留, 则 $\sum\limits_{n=1}^{\infty} p_{x,y}^{(n)} < \infty$, $x \in E$.
6. 设 $E = \{0, 1\}$.

   (1) 验证: $p_{0,0}^{(n)} = p_{1,0} + (p_{0,0} - p_{1,0}) p_{0,0}^{(n-1)}$;

   (2) 写出 $p_{0,0}^{(n)}$ 的表达式并求出极限;

   (3) 求出 $p_{x,y}^{(n)}$ 及其极限.
7. 证明: (1) 如果 C 是闭的, 那么对任何 $n \geqslant 1$, $x \in C$, $y \notin C$ 有 $p_{x,y}^{(n)} = 0$;

   (2) 一个 Markov 链不可分当且仅当它没有真闭子集;

   (3) 所有常返状态全体是闭集.

8. 甲袋中有 3 个黑球, 乙袋中有 3 个白球, 每次从两袋中各任取一球, 交换放入另一袋中, 用 p_n 表示交换 n 次后甲袋中的球颜色一致的概率, 求 $\lim_n p_n$.

9. 设 $X = (X_n)$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上以 E 为状态空间的 Markov 链. 令 $Y_n := (X_n, X_{n+1})$. 记 $F := \{(x, y) \in E \times E:\ p_{x,y} > 0\}$.

   (1) 证明: (Y_n) 是一个以 F 为状态空间的 Markov 链;

   (2) 写出转移概率并证明: 如果 X 是不可分与非周期, 则 Y 也是;

   (3) 证明: 如果 $\{\pi_x\}$ 是 X 的平稳分布, 则 $\{\pi_x p_{x,y} : (x, y) \in F\}$ 是 Y 的平稳分布.

10. 证明: 一个有限状态 Markov 链没有零常返状态, 且不能所有状态都是暂留的.

11. 某人有 r 把伞放在家或者办公室用于来往于家与办公室之间. 当且仅当天下雨且手边有伞时, 带一把伞走, 到达后放下. 下雨的概率等于 p.

   (1) 用 X_n 表示他第 n 次出(家或者办公室)门时手边的伞的数目, 说明它是一个 Markov 链, 写出其转移概率. 它是否有平稳分布? 如有, 求其平稳分布.

   (2) 计算他被淋湿的概率的极限.

   (3) 证明: 对任何 p, 5 把伞可以保证他以 95% 以上的概率不被淋湿.

12. E 上的函数 h 称为关于 Markov 链 X 不变的, 如果对任何 $x \in E$, $h(x) = \sum\limits_{y \in E} p_{x,y} h(y)$.

   (1) 验证: 对任何 $a, b \in E$,$x \mapsto \mathbb{P}^x(\tau_a < \tau_b)$ 是调和函数;

   (2) 证明: 如果 X 不可分,那么某点处为零的非负不变函数恒等于零.

13. 设 C 是闭集, $x \notin C$. 证明: 如果 x 常返, 则 x 不可达 C 中的任何状态. 因此常返的 Markov 链的状态空间可以唯一分解为闭集的不交并, 使得 X 限制在其中每个闭集上不可分.

*14. 设 X 是常返的 Markov 链, 那么 X 一定有非平凡的不变测度. 即存在 E 上的非零测度 $(\pi_x)_{x \in E}$, 使得 $\sum\limits_{x \in E} \pi_x p_{x,y} = \pi_y$ 对任何 $y \in E$ 成立. 如果 X 是不可分的, 那么 π 在相差常数倍的意义下唯一.

§2.5 Poisson过程

一个随机过程也被视为从概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 到 $(E^{\mathsf{T}},\mathscr{E}^{\mathsf{T}})$ 的一个可测映射 ξ, 即对 $\omega\in\Omega$, $\xi\omega: t\mapsto X_t(\omega)$ 是 T 到 E 的映射, 因此随机过程也常称为是随机映射或随机函数.

现在我们设 T 上有一个可测结构, 即 σ- 代数 $\mathscr{E}(\mathsf{T})$, 用 M 表示 $(\mathsf{T},\mathscr{E}(\mathsf{T}))$ 上测度全体, 且对于 $B\in\mathscr{E}(\mathsf{T})$, $\mu\mapsto\mu(B)$ 是 M 上(可取无穷大值)的非负函数, 令

$$\mathscr{E}_M := \sigma(\{\mu\mapsto\mu(B): B\in\mathscr{E}(\mathsf{T})\}),$$

即是 M 上使函数 $\mu\mapsto\mu(B)$, $B\in\mathscr{E}(\mathsf{T})$ 成为可测函数的最小 σ- 代数. 概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 到 $(M,\mathscr{E}_M)$ 的一个可测映射称为是 T 上的一个随机测度. 显然 T 上的随机测度 μ 自然地诱导了 $(\Omega,\mathscr{F})$ 到 $(\mathsf{T},\mathscr{E}(\mathsf{T}))$ 的核: $\mu(\omega,B):=\mu(\omega)(B)$, 反之 $(\Omega,\mathscr{F})$ 到 $(\mathsf{T},\mathscr{E}(\mathsf{T}))$ 的核也可自然地视为 T 上的随机测度. 另外一个随机测度也可视为一个指标集为 $\mathscr{E}(\mathsf{T})$ 的随机过程. 如果 μ 是随机测度, 定义 $\lambda(B):=E\mu(\cdot,B)$, $B\in\mathscr{E}(\mathsf{T})$. 由 Fubini 定理, λ 是 $(\mathsf{T},\mathscr{E}(\mathsf{T}))$ 上的测度, 称为是 μ 的强度.

在这一节中, 我们将仅研究整数值随机测度, 实际上, 点过程是整数值随机测度的特例. 用 M_0 表示 $(\mathsf{T},\mathscr{E}(\mathsf{T}))$ 上的非负整数值 (可取无穷大)测度全体, $\mathscr{E}_{M_0}:=M_0\cap\mathscr{E}_M$, 即 M_0 上使函数 $\mu\mapsto\mu(B)$, $B\in\mathscr{E}(\mathsf{T})$ 成为可测函数的最小 σ-代数. 概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 到 $(M,\mathscr{E}_M)$ 的一个可测映射称为是 T 上的一个整值随机测度, 我们通常以核的形式来记一个随机测度.

定义 2.5.1 设 $(\Omega,\mathscr{F},\mathbb{P})$ 是概率空间, $(\mathsf{T},\mathscr{E}(\mathsf{T}))$ 是可测空间, T 上的一个整值随机测度 μ 称为是 Poisson 随机测度, 如果下列条件满足:

(1) 对 $B\in\mathscr{E}(\mathsf{T})$, $\lambda(B)<\infty$, 则 $\mu(\cdot,B)$ 服从参数为 $\lambda(B)$ 的 Poisson 分布;

(2) 若 $B_1,\cdots,B_n\in\mathscr{E}(\mathsf{T})$ 互不相交且 $\lambda(B_i)<\infty$, $1\leqslant i\leqslant n$, 则随机变量

$$\mu(\cdot,B_1),\cdots,\mu(\cdot,B_n)$$

是相互独立的.

下面是 Poisson 随机测度的存在性定理.

定理 2.5.1 给定 $(\mathsf{T},\mathscr{E}(\mathsf{T}))$ 上的一个 σ- 有限测度 λ, 存在一个强度为 λ 的 Poisson 随机测度 μ.

证明 因 λ 是 σ- 有限的, 故存在互不相交的集列 $\{U_n\}\subset\mathscr{E}(\mathsf{T})$, 使得 $0<\lambda(U_n)<\infty$ 且 $\bigcup_n U_n=\mathsf{T}$, 则由推论 2.2.1, 存在一个概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 及其上满足如下条件的随机变量族:

(1) 对任何 $n\geqslant 1$, $\{\xi_i^{(n)}:\ i=1,2,\cdots\}$ 是一列取值于 U_n 的随机变量, 且分布同是 $\dfrac{1}{\lambda(U_n)}\cdot\lambda$;

(2) 对任何 $n\geqslant 1$, s_n 是服从参数为 $\lambda(U_n)$ 的 Poisson 分布的随机变量;

(3) 所有随机变量 $\xi_i^{(n)},s_n,\ n,i=1,2,\cdots$ 是相互独立的.

令

$$\mu(\omega,B):=\sum_{n=1}^{\infty}\sum_{i=1}^{s_n(\omega)}1_{B\cap U_n}(\xi_i^{(n)}(\omega)),\ \omega\in\Omega,\ B\in\mathscr{E}(\mathsf{T}),$$

(和式中如出现 $\sum\limits_{i=1}^{0}$ 约定是零.) 则 μ 是一个满足定理要求的 Poisson 随机测度. 下面我们对此进行证明.

任取 $B_1,\cdots,B_m\in\mathscr{E}(\mathsf{T})$ 互不相交及非负实数 $a_1,\cdots,a_m$. 我们需要计算随机向量 $(\mu(\cdot,B_1),\cdots,\mu(\cdot,B_m))$ 的 Laplace 变换

$$(a_1,\cdots,a_m)\mapsto\mathbb{E}\exp\left(-\sum_{j=1}^{m}a_j\mu(\cdot,B_j)\right).$$

首先, 记 $B_j^n:=B_j\cap U_n$. 对 $n\geqslant 1$, 令

$$X_n:=\sum_{i=1}^{s_n}\sum_{j=1}^{m}a_j1_{B_j^n}(\xi_i^{(n)}),$$

则 $\{X_n\}$ 是独立随机变量序列.

事实上, 任取 $\mathbf{R}$ 的 Borel 子集 $A_1,\cdots,A_n$,有

$$\mathbb{P}(X_1\in A_1,\cdots,X_n\in A_n)$$

$$= \sum_{j_1,\cdots,j_n \geqslant 0} \mathbb{P}(X_1 \in A_1, \cdots, X_n \in A_n, s_1 = j_1, \cdots, s_n = j_n)$$

$$= \sum_{j_1,\cdots,j_n \geqslant 0} \mathbb{P}\left(\sum_{i=1}^{j_k}\sum_{j=1}^{m} a_j 1_{B_j^k}(\xi_i^{(k)}) \in A_k, s_k = j_k, 1 \leqslant k \leqslant n\right)$$

$$= \sum_{j_1,\cdots,j_n \geqslant 0} \prod_{k=1}^{n} \mathbb{P}\left(\sum_{i=1}^{j_k}\sum_{j=1}^{m} a_j 1_{B_j^k}(\xi_i^{(k)}) \in A_k, s_k = j_k\right)$$

$$= \prod_{k=1}^{n} \sum_{j_k \geqslant 0} \mathbb{P}\left(\sum_{i=1}^{j_k}\sum_{j=1}^{m} a_j 1_{B_j^k}(\xi_i^{(k)}) \in A_k, s_k = j_k\right)$$

$$= \prod_{k=1}^{n} \mathbb{P}\left(\sum_{i=1}^{s_k}\sum_{j=1}^{m} a_j 1_{B_j^k}(\xi_i^{(k)}) \in A_k\right).$$

现在

$$\mathbb{E}\exp\left(-\sum_{j=1}^{m} a_j \mu(\cdot, B_j)\right)$$

$$= \mathbb{E}\exp\left(-\sum_{n=1}^{\infty}\sum_{i=1}^{s_n}\sum_{j=1}^{m} a_j 1_{B_j^n}(\xi_i^{(n)})\right)$$

$$= \prod_{n=1}^{\infty} \mathbb{E}\exp\left(-\sum_{i=1}^{s_n}\sum_{j=1}^{m} a_j 1_{B_j^n}(\xi_i^{(n)})\right)$$

$$= \prod_{n=1}^{\infty}\sum_{l=0}^{\infty} \mathbb{E}\exp\left(-\sum_{i=1}^{l}\sum_{j=1}^{m} a_j 1_{B_j^n}(\xi_i^{(n)})\right)\mathbb{P}(s_n = l)$$

$$= \prod_{n=1}^{\infty}\sum_{l=0}^{\infty} \mathbb{P}(s_n = l)\prod_{i=1}^{l} \mathbb{E}\exp\left(-\sum_{j=1}^{m} a_j 1_{B_j^n}(\xi_i^{(n)})\right)$$

$$= \prod_{n=1}^{\infty}\sum_{l=0}^{\infty} \mathbb{P}(s_n = l)\left(\int_{U_n} \frac{\exp\{-\sum_{j=1}^{m} a_j 1_{B_j \cap U_n}(u)\}}{\lambda(U_n)}\lambda(\mathrm{d}u)\right)^l$$

$$= \prod_{n=1}^{\infty} \sum_{l=0}^{\infty} \mathbb{P}(s_n = l) \left(1 - \frac{\sum_{j=1}^{m} (1 - \mathrm{e}^{-a_j}) \lambda(B_j \cap U_n)}{\lambda(U_n)} \right)^l$$

$$= \prod_{n=1}^{\infty} \mathrm{e}^{-\lambda(U_n)} \sum_{l=0}^{\infty} \frac{1}{l!} \left(\lambda(U_n) - \sum_{j=1}^{m} (1 - \mathrm{e}^{-a_j}) \lambda(B_j \cap U_n) \right)^l$$

$$= \prod_{n=1}^{\infty} \exp \left(-\lambda(U_n) + \lambda(U_n) - \sum_{j=1}^{m} (1 - \mathrm{e}^{-a_j}) \lambda(B_j \cap U_n) \right)$$

$$= \exp \left\{ - \sum_{n=1}^{\infty} \sum_{j=1}^{m} (1 - \mathrm{e}^{-a_i}) \lambda(B_j \cap U_n) \right\}$$

$$= \exp \left\{ - \sum_{j=1}^{m} (1 - \mathrm{e}^{-a_j}) \lambda(B_j) \right\}.$$

这个等式最终证明了两件事情: (i) 对任何 $B \in \mathscr{E}(\mathsf{T})$, $\mu(\cdot, B)$ 的 Laplace 变换是

$$\mathbb{E} \exp(-a\mu(\cdot, B)) = \exp(-(1 - \mathrm{e}^{-a})\lambda(B)),$$

因此 $\mu(\cdot, B)$ 服从参数为 $\lambda(B)$ 的 Poisson 分布; (ii) $\mu(\cdot, B_1), \cdots, \mu(\cdot, B_m)$ 是相互独立的. □

取 $\mathsf{T} = [0, \infty)$, $\mathscr{E}(\mathsf{T})$ 是 T 上的 Borel 代数, m 是 T 上的 Lebesgue 测度, $\lambda > 0$ 是常数, 则存在 T 上的随机测度 μ 使得对 $B \in \mathscr{E}(\mathsf{T})$, $\mathbb{E}(\mu(\cdot, B)) = \lambda m(B)$. 对 $t \in \mathsf{T}$, $\omega \in \Omega$, 令

$$N_t(\omega) := \mu(\omega, (0, t]),$$

则对于 $t > s > 0$, 有 $N_t - N_s = \mu(\cdot, (s, t])$, 且过程 $N = (N_t : t \in \mathsf{T})$ 有下列性质:

(1) $N_0 = 0$, $t \mapsto N_t(\omega)$, $\omega \in \Omega$ 是单调增加的右连续整数值阶梯函数;

(2) 对 $t > s \geqslant 0$, $N_t - N_s$ 服从参数为 $\lambda(t - s)$ 的 Poisson 分布;

(3) N 是独立增量过程, 即对于 $0 \leqslant t_1 < \cdots < t_n$, $N_{t_1}, N_{t_2} - N_{t_1}, \cdots, N_{t_n} - N_{t_{n-1}}$ 是相互独立的.

定义 2.5.2 定义在带流概率空间 $(\Omega,\mathscr{F},(\mathscr{F}_t),\mathbb{P})$ 上的适应右连续随机过程 $N=(N_t:t\in\mathsf{T})$ 称为是参数为 λ 的 $(\mathscr{F}_t)$-Poisson 过程, 如果它满足:

(1) $N_0=0$ a.s.;

(2) 对 $t>s\geqslant 0$, $k\geqslant 0$,

$$\mathbb{P}(N_t-N_s=k|\mathscr{F}_s)=\frac{(\lambda(t-s))^k}{k!}\mathrm{e}^{-\lambda(t-s)}.$$

Poisson 过程是 Poisson 随机测度的一个特例, 上面由 Poisson 随机测度构造的满足性质 (1)(2)(3)的过程是关于自然流的 Poisson 过程, 因此 Poisson 随机测度的存在性蕴含着 Poisson 过程的存在性.

设 $N=(N_t:t\in\mathsf{T})$ 是 Poisson 过程, 则它在 t 时刻的值 N_t 常被视为一个服务系统在 t 时刻前请求服务的次数, 令

$$S_n:=\inf\{t\in\mathsf{T}:N_t\geqslant n\},\ n\geqslant 0. \tag{2.5.1}$$

S_n 理解为第 n 次服务请求的时刻, 再令 T_n 是服务请求间隔时间, $T_n:=S_n-S_{n-1}$, $n\geqslant 1$.

定理 2.5.2 $\{T_n:n\geqslant 1\}$ 是独立同分布的, 且都服从参数为 λ 的指数分布.

证明 先计算 $(S_1,S_2,\cdots,S_k)$ 的联合分布, 显然它的值域是

$$G:=\{(y_1,\cdots,y_k)\in\mathbf{R}^k:0<y_1<\cdots<y_k\}.$$

考虑矩形 $(s_1,t_1]\times(s_2,t_2]\times\cdots\times(s_k,t_k]\subset G$, 则 $0<s_1<t_1<s_2<t_2<\cdots<s_k<t_k$, 且

$$\begin{aligned}
&\mathbb{P}((S_1,\cdots,S_k)\in(s_1,t_1]\times\cdots\times(s_k,t_k])\\
&=\mathbb{P}(s_1<S_1\leqslant t_1<s_2<S_2\leqslant t_2<\cdots<s_k<S_k\leqslant t_k)\\
&=\mathbb{P}(N_{s_1}=0,N_{t_1}-N_{s_1}=1,N_{s_2}-N_{t_1}=0,\\
&\quad\cdots,N_{s_k}-N_{t_{k-1}}=0,N_{t_k}-N_{s_k}\geqslant 1)\\
&=\mathbb{P}(N_{s_1}=0)\mathbb{P}(N_{t_1}-N_{s_1}=1)\\
&\quad\cdots\mathbb{P}(N_{s_k}-N_{t_{k-1}}=0)\mathbb{P}(N_{t_k}-N_{s_k}\geqslant 1)
\end{aligned}$$

$$= e^{-\lambda(s_1+s_2-t_1+\cdots+s_k-t_{k-1})}\cdot(t_1-s_1)\cdots(t_{k-1}-s_{k-1})$$

$$\cdot\lambda^{k-1}e^{-\lambda(t_1-s_1+\cdots+t_{k-1}-s_{k-1})}(1-e^{-\lambda(t_k-s_k)})$$

$$=\lambda^{k-1}(t_1-s_1)\cdots(t_{k-1}-s_{k-1})(e^{-\lambda s_k}-e^{-\lambda t_k})$$

$$=\int_{(y_1,\cdots,y_k)\in(s_1,t_1]\times\cdots\times(s_k,t_k]}\lambda^k e^{-\lambda y_k}dy_1\cdots dy_k,$$

另外 G 中的矩形生成的 σ- 代数恰是 G 上的 Borel 代数, 故 k-维随机向量 $(S_1,S_2,\cdots,S_k)$ 的分布密度函数是

$$p(y_1,\cdots,y_k)=\lambda^k e^{-\lambda y_k}1_{\{0\leqslant y_1<\cdots<y_k\}}.$$

现在我们来计算 $(T_1,\cdots,T_k)$ 的联合分布. 对于 $t_1,\cdots,t_k>0$,

$$\mathbb{P}(T_1>t_1,\cdots,T_k>t_k)$$

$$=\mathbb{P}(S_1>t_1,S_2-S_1>t_2,\cdots,S_n-S_{n-1}>t_n)$$

$$=\int_{y_1>t_1,y_2-y_1>t_2,\cdots,y_k-y_{k-1}>t_k}\lambda^k e^{-\lambda y_k}dy_1\cdots dy_k$$

$$=\int_{y_1>t_1,y_2-y_1>t_2,\cdots,y_{k-1}-y_{k-2}>t_{k-1}}\lambda^{k-1}\int_{t_k+y_{k-1}}^{\infty}(-de^{\lambda y_k})$$

$$=\cdots=e^{-\lambda t_1}e^{-\lambda t_2}\cdots e^{-\lambda t_k}.$$

这证明了首先每个 T_n 都服从参数为 λ 的指数分布, 其次 $\{T_n\}$ 是相互独立的. □

反过来, 给定概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 上的一个服从参数为 λ 的指数分布的独立随机变量列 $\{T_n:\ n\geqslant 1\}$. 令

$$S_0:=0,\ S_n:=\sum_{i=1}^n T_i,$$

则不难验证 $(S_1,S_2,\cdots,S_k)$ 的联合分布密度是

$$p(y_1,\cdots,y_k)=\lambda^k e^{-\lambda y_k}1_{\{0\leqslant y_1<\cdots<y_k\}}.$$

因此任取整数 $1\leqslant k_1<k_2<\cdots<k_n$, $(S_{k_1},S_{k_2},\cdots,S_{k_n})$ 的联合分布密度是

$$\frac{y_1^{k_1-1}}{(k_1-1)!}\frac{(y_2-y_1)^{k_2-k_1-1}}{(k_2-k_1-1)!}\cdots\frac{(y_n-y_{n-1})^{k_n-k_{n-1}-1}}{(k_n-k_{n-1}-1)!}\lambda^{k_n}\mathrm{e}^{-\lambda y_n},$$

$0<y_1<y_2<\cdots<y_n$. 令

$$N_t:=\sup\{n\geqslant 0: S_n\leqslant t\},\ t\geqslant 0,$$

即 $N_t=k$ 当且仅当 $S_k\leqslant t<S_{k+1}$.

定理 2.5.3　以上构造的随机过程 $N=(N_t:t\in\mathsf{T})$ 是参数为 λ 的关于自然流的 Poisson 过程.

证明　Poisson 过程定义的条件 (1) 是显然的. 对于 $0\leqslant t_1<\cdots<t_n$, 应用上面给出的联合分布密度计算 $N_{t_1},N_{t_2}-N_{t_1},\cdots,N_{t_n}-N_{t_{n-1}}$ 的联合分布律, 得

$$\begin{aligned}&\mathbb{P}(N_{t_1}=j_1,N_{t_2}-N_{t_1}=j_2,\cdots,N_{t_n}-N_{t_{n-1}}=j_n)\\=&\mathbb{P}(N_{t_1}=j_1,N_{t_2}=j_1+j_2,\cdots,N_{t_n}=j_1+\cdots+j_n)\\=&\mathrm{e}^{-\lambda t_1}\frac{(\lambda t_1)^{j_1}}{j_1!}\cdot\mathrm{e}^{-\lambda(t_2-t_1)}\frac{(\lambda(t_2-t_1))^{j_2}}{j_2!}\cdots\mathrm{e}^{-\lambda(t_n-t_{n-1})}\frac{(\lambda(t_n-t_{n-1}))^{j_n}}{j_n!},\end{aligned}$$

其中 $j_1,\cdots,j_n$ 是任何非负整数, $k_i:=j_1+\cdots+j_i$, $i=1,\cdots,n$. 这表示 N 是关于自然流的 Poisson 过程. □

Poisson 过程的构造还有多种方法, 例如由第四节的结论, 存在一个从零点出发的 Markov 过程 X 具有 Poisson 半群作为转移半群, 过程 X 是一个平稳独立增量过程且增量服从要求的 Poisson 分布, 但 Kolmogorov 定理不保证轨道有所要求的性质. 我们可以用类似下一节中证明 Brown 运动轨道连续时所使用的方法证明 X 有一个修正是 Poisson 过程.

最后我们介绍点过程的概念, 可以说它是随机测度的特例, 但在实际中有更多的应用. 设 $(E,\mathscr{E})$ 是可测空间, 给 E 附加一个点 $\delta\notin E$ 记为 $E_\delta,\mathscr{E}_\delta$ 是相应的 σ- 代数.

定义 2.5.3　概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 上, 状态空间为 $(E_\delta,\mathscr{E}_\delta)$ 的过程 $p=(p_t:t>0)$ 称为点过程, 如果

(1) 映射 $(s,\omega)\mapsto p_t(\omega)$ 是 $\mathscr{B}((0,\infty))\times\mathscr{F}$ 可测的;

(2) 对几乎所有 ω, $D(\omega):=\{t:\ p_t(\omega)\neq\delta\}$ 是可数的.

点过程 p 称为是平稳的, 如果对任何 $t>0$, p 与 $p\circ\theta_t$ 等价, 其中 $p\circ\theta_t$ 是指点过程 $s\mapsto p_{s+t}$.

设 p 是点过程. 由定义

$$N((0,t]\times U):=\#\{s:\ 0<s\leqslant t,p(s)\in U\},\ t>0,U\in\mathscr{E}$$

所诱导的 $(0,\infty)\times E$ 的随机测度 N 称为是 p 对应的随机测度. 点过程 p 称为离散的, 如果对任何 $t>0$, $N((0,t]\times E)<\infty$ a.s.; 称为是 σ- 离散的, 如果存在 $E_n\uparrow E$, 使得 p 限制在每个 E_n 上是离散的. 一个 σ- 离散的点过程 p 称为是 Poisson 的, 如果 N 是 Poisson 随机测度. 显然, 如果 p 是平稳 Poisson 点过程, 那么 N 的强度 $n(\mathrm{d}t\mathrm{d}x)$ 关于 t 平移不变, 即存在 E 上的测度 n, 使得 $n(\mathrm{d}t\mathrm{d}x)=\mathrm{d}t\,n(\mathrm{d}x)$. 这时 n 显然是一个 σ- 有限测度, 称为 p 的特征测度. 由定理 2.5.1 推出下面平稳 Poisson 点过程的存在定理.

定理 2.5.4　给定 $\mathbb{E}$ 上 σ- 有限测度 n, 存在一个特征测度为 n 的 $\mathbb{E}$ 上的平稳 Poisson 点过程.

如果 $N=(N_t)$ 是 Poisson 过程, 那么 $p_t:=\Delta N_t:=N_t-N_{t-}$ 是 $\mathbf{R}\backslash\{0\}$ 上的平稳 Poisson 点过程, 特征测度是测度 $\lambda\epsilon_1$. 反之亦然. 下面我们将给出平稳 Poisson 点过程的一个刻画. 先证明一个引理. 一个适应过程 $N=(N^1,\cdots,N^d)$ 称为是 d-维 Poisson 过程, 如果每个 N^i 都是右连续适应过程, $N_0^i=0$ 且存在 λ_i, 使得对任何 $t>s>0$, $k_1,\cdots,k_d\in\mathbf{Z}_+$,

$$\mathbb{P}\left(\bigcap_{i=1}^{d}\{N_t^i-N_s^i=k_i\}\big|\mathscr{F}_s\right)=\prod_{i=1}^{d}\frac{(\lambda_i(t-s))^{k_i}}{k_i!}\mathrm{e}^{-\lambda_i(t-s)}.$$

引理 2.5.1　一个适应过程 $N=(N^1,\cdots,N^d)$ 是 d-维 Poisson 过程当且仅当

(1) 每个 N^i 是 Poisson 过程;

(2) N 是平稳独立增量过程;

(3) 任何两个 N^i 都不会同时跳跃.

证明　我们只证明充分性, 必要性作为习题. 不妨设 $d=2$. 在条件之下, 只需验证 N^1,N^2 是独立的. N 诱导 $\mathbf{R}^2\backslash\{0\}$ 上的点过程 $p_t:=(\Delta N_t^1,\Delta N_t^2)$. 条件 (1), (2) 说明 p 是离散的平稳 Poisson 点过程. 因为 N^1,N^2 不会同时跳跃, 故

$$N_t^1 = \#\{0 < s \leqslant t:\ p_s = (1,0)\}$$

$$N_t^2 = \#\{0 < s \leqslant t:\ p_s = (0,1)\}.$$

由定义推出两者是独立的. □

定理 2.5.5 一个 σ- 离散的点过程 p 是平稳 Poisson 点过程当且仅当对任何 $s > 0, 0 < t_1 < \cdots < t_d$ 及 $U_i \in \mathscr{E}$, 使得 $(N((s, s+t_i] \times U_i) : 1 \leqslant i \leqslant d)$ 独立于 $\mathscr{F}_s$ 且与 $(N((0, t_i] \times U_i) : 1 \leqslant i \leqslant d)$ 同分布, 其中 $(\mathscr{F}_t)$ 是 p 的自然流, N 是对应的随机测度.

证明 只需证明充分性. 在给定的条件下容易看出 p 是平稳的. 我们只需证明 N 是 Poisson 随机测度. 首先由条件推出过程 $t \mapsto N((0,t] \times U)$ 是平稳独立增量右连续过程且以跃度 1 增加, 因此是一个 Poisson 过程(参考定理 4.3.9), 那么我们只需验证随机测度定义的条件 (2). 再由条件所示的独立性, 我们只需证明对任何 $U_i : i = 1, \cdots, d$ 互不相交, $(N((0,t] \times U_i) : i = 1, \cdots, d)$ 独立就足够了. 而这恰是 d 个具平稳独立增量的且任何两个都不会同时跳跃的 Poisson 过程. 引理 2.5.1 说明它们是独立的. □

例 2.5.1 如果 f 是 $[0, \infty)$ 上右连续左极限存在的函数, 那么容易验证 f 的不连续点是至多可数的. 实际上, 它在任何有限区间上跃度不小于任何给定正数的不连续点 $\{t : |f(t) - f(t-)| > a\}$, $a > 0$ 是有限的. 设 X 是 E 上的一个右连续且左极限存在的随机过程, 那么 X 的几乎所有轨道的不连续点最多是可数的. 定义

$$p_t := (X_{t-}, X_t),\ t > 0,$$

那么 p 是 $E \times E \backslash \delta$ 上的 σ- 离散的点过程, 其中 δ 是 $E \times E$ 的对角线.

如果 X 是 $\mathbf{R}^d$ 上一个右连左极的平稳独立增量过程(参考例 2.1.6), 那么令 $p_t = \Delta X_t = X_t - X_{t-}$, $t > 0$. p 是一个 $\mathbf{R}^d \backslash \{0\}$ 上的平稳 Poisson 点过程. 事实上, 因为若 s 固定,

$$\begin{aligned} N((s, s+t] \times U) &= \sum_{s < r \leqslant s+t} 1_{\{\Delta X_r \in U\}} \\ &= (\sum_{0 < r \leqslant t} 1_{\{\Delta X_r \in U\}}) \circ \theta_s = N((0,t] \times U) \circ \theta_s, \end{aligned}$$

由过程 $t \mapsto X_{t+s} - X_s$ 完全决定, 故由平稳独立增量的性质推出对任何 $s > 0$, $0 < t_1 < \cdots < t_d$ 及 $U_i \in \mathscr{B}(\mathbf{R}^d \backslash \{0\})$, $(N((s, s+t_i] \times U_i) : 1 \leqslant i \leqslant d)$ 独立于 $\mathscr{F}_s$ 且与 $(N((0, t_i] \times U_i) : 1 \leqslant i \leqslant d)$ 同分布. 因此由定理 2.5.5 推出 p 是平稳 Poisson 点过程. 后面我们将证明其特征测度就是 X 的 Lévy 测度. ▮

习　题

1. 设 $X_i : i \geqslant 1$ 是独立随机序列, 分布都是 μ, N 服从参数为 $\lambda > 0$ 的 Poisson 分布, 且与 $\{X_i\}$ 独立, 定义 $\eta(B) := \sum_{i=1}^{N} 1_{\{X_i \in B\}}$, 证明: η 是 Poisson 随机测度.

2. 设 N 是参数为 λ 的 Poisson 过程, S_n 如(2.5.1)定义. 固定 $t \geqslant 0$. 令 $A_t := t - S_{N_t}$, $B_t = S_{N_t+1} - t$. 证明: A_t 与 B_t 独立. 试求它们的分布.

3. S_n 如上. 固定 $t \geqslant 0$. 定义 $L_t := S_{N_t+1} - S_{N_t}$. 证明: L_t 的密度函数为
$$d_t(x) = \begin{cases} \lambda^2 x \mathrm{e}^{-\lambda x} & x \in (0, t), \\ \alpha(1 + \lambda t)\mathrm{e}^{-\lambda x} & x \geqslant t. \end{cases}$$

4. 设有 Poisson 过程 $N = (N_t : t \geqslant 0)$ 与 $(X_n : n \geqslant 1)$ 是独立同分布随机序列, 且 $\sigma(X_n : n \geqslant 1)$ 独立于 $\sigma(N_t : t \geqslant 0)$. 令 $Z_t := \sum_{k \leqslant N_t} X_k$. 证明: $Z = (Z_t : t \geqslant 0)$ 是一个右连续的平稳独立增量过程, 称为复合 Poisson 过程.

5. 设 $0 \leqslant s < t$, 证明:
$$\mathbb{P}(N_s = k | N_t = n) = \binom{n}{k} \left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{n-k}, \ 0 \leqslant k \leqslant n.$$

6. 设 $N = (N_t : t \geqslant 0)$ 与 $N' = (N'_t : t \geqslant 0)$ 是两个独立且参数分别为 λ 与 λ' 的 Poisson 过程. 证明:

   (1) $(N_t + N'_t)$ 是参数为 $\lambda + \lambda'$ 的 Poisson 过程;

   (2) $(N_t + N'_t)$ 的第一次跳跃来自过程 N 的概率是 $\dfrac{\lambda}{\lambda + \lambda'}$.

7. 设一部机器需两种部件才能运作, 1, 2 两种部件的库存分别为 n, m 个, 且每个部件的工作时间独立分别服从参数为 α_1, α_2 的指数分布. 计算机器可运作时间的平均长度.

8. 电梯从 1 层出发向上, N_i 表示第 i 层上电梯的人数, 它们相互独立并服从参数为 λ_i 的 Poisson 分布. 独立于其他事件, 从 i 层上电梯的每个人在 j 层下的概率是 p_{ij}, $\sum\limits_{j>i} p_{ij} = 1$. 再用 O_j 表示 j 层下电梯的人数.

   (1) 计算 $\mathbb{E}(O_j)$;

   (2) 计算 O_j 的分布;

   (3) 计算 O_j, O_k 的联合分布.

9. 计算给定 $S_n = t$ 时, $(S_1, \cdots, S_{n-1})$ 的条件联合密度.

10. 设 m 是 $\mathbf{R}^2$ 上 Lebesgue 测度. μ 是以 λm 为特征测度的 Poisson 随机测度. 记 $X(\omega)$ 是 0 点到测度 $\mu(\omega, \cdot)$ 的最近的支撑点的距离. 证明: $\mathbb{P}(X > t) = \mathrm{e}^{-\lambda\pi t^2}$ 且 $\mathbb{E}X = \dfrac{1}{2}\sqrt{\lambda}$.

11. 设 N^1, N^2 是两个独立的 Poisson 过程, 证明: 它们几乎不可能同时跳跃, 即

$$\mathbb{E}\sum_{s>0}(N_s^1 - N_{s-}^1)(N_s^2 - N_{s-}^2) = 0.$$

12. 设 X, Y 是两个等价的右连左极平稳独立增量过程. 证明: 它们对应的平稳 Poisson 点过程的特征测度一致.

§2.6 Brown运动

Brown 运动是最重要的一类随机过程. 可以说, 它是所有随机过程重要性质的交集. 如它是 Markov 过程, 它是平稳独立增量的, 它是连续的, 它是鞅等等. Brown 运动实际上是对植物学家 R.Brown 所发现现象的一种数学模拟. 他发现悬浮在液体表面的花粉颗粒会无序地向各个方向运动而无法预测. 在 1905 年, A. Einstein 用自己非凡的物理直觉找到了答案: 花粉颗粒的无序运动是因为水分子在各个方向撞击它们的缘故, 从而算出了它的增量的分布. N.Wiener 在 20 年后证明了按这种分布得到的随机过程的轨道实际上是连续的, 因此它的确可以认为是花粉颗粒运动的数学描述. 所以 Brown 运动也称为 Wiener 过程.

考虑 $\mathbf{R}^d$ 上的热传导方程

$$\frac{\partial u}{\partial t} = \frac{1}{2}\triangle u,$$

其中 $\triangle$ 是 Laplace 算子. 容易验证方程的解是

$$p(t,x) := \frac{1}{(2\pi t)^{\frac{d}{2}}}\mathrm{e}^{-\frac{|x|^2}{2t}},\ t>0, x\in \mathbf{R}^d.$$

对 $t>0$, 令 $P_t(x,A) := \int_A p(t,y-x)\mathrm{d}y,\ x\in\mathbf{R}^d, A\in\mathscr{B}(\mathbf{R}^d)$, 称其为热核半群(heat kernel).

引理 2.6.1　热核半群 (P_t) 是 $\mathbf{R}^d$ 上转移半群且满足

(1) 对任何有界的 $f\in\mathscr{B}(\mathbf{R}^d)$ 及 $t>0$, $P_t f$ 是有界连续的;

(2) 对 $f\in C_0(\mathbf{R}^d)$, 当 $t\downarrow 0$ 时, $P_t f$ 一致收敛于 f.

证明　先证明 (P_t) 是转移半群. 不妨设 $d=1$, 首先对任何 $t>0$, $P_t 1=1$, 其次对 $t,s>0$,

$$\begin{aligned}
P_tP_s(x,A) &= \int_{\mathbf{R}} P_t(x,\mathrm{d}y)P_s(y,A)\\
&= \int_A \mathrm{d}z\int_{\mathbf{R}} p(t,y-x)p(s,z-y)\mathrm{d}y\\
&= \int_A \mathrm{d}z\int_{\mathbf{R}} \frac{1}{2\pi\sqrt{ts}}\exp\left(-\frac{(y-x)^2}{2t}-\frac{(z-y)^2}{2s}\right)\mathrm{d}y\\
&= \int_A \mathrm{d}z\int_{\mathbf{R}} \frac{1}{2\pi\sqrt{ts}}\exp\left(-\frac{(s+t)y^2-2(xs+zt)y+tz^2+sx^2}{2ts}\right)\\
&= \int_A \mathrm{d}z\int_{\mathbf{R}} \frac{1}{2\pi\sqrt{ts}}\exp\left(-\frac{s+t}{2st}\left(y-\frac{sx+zt}{t+s}\right)^2\right)\exp\left(-\frac{(z-x)^2}{2(t+s)}\right)\mathrm{d}y\\
&= \int_A \mathrm{e}^{-\frac{(z-x)^2}{2(t+s)}}\frac{1}{2\pi\sqrt{ts}}\cdot\sqrt{2\pi\frac{ts}{t+s}}\mathrm{d}z = P_{t+s}(x,A).
\end{aligned}$$

另外因为 $p(t,\cdot)$ 是连续的, 故 (1) 显然. 最后证明 (2), 取 $f\in C_0(\mathbf{R}^d)$, 那么 f 是一致连续的且对任何 $N>0$, 有

$$
\begin{aligned}
|P_t f(x) - f(x)| &\leqslant \int_{\mathbf{R}} |f(y+x) - f(x)| p(t,y) \mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} |f(\sqrt{t}y + x) - f(x)| \mathrm{e}^{-\frac{|y|^2}{2}} \mathrm{d}y \\
&\leqslant \frac{1}{\sqrt{2\pi}} \int_{|y|\leqslant N} |f(\sqrt{t}y + x) - f(x)| \mathrm{e}^{-\frac{|y|^2}{2}} \mathrm{d}y \\
&\quad + \frac{2\|f\|}{\sqrt{2\pi}} \int_{|y|>N} \mathrm{e}^{-\frac{|y|^2}{2}} \mathrm{d}y.
\end{aligned}
$$

因此由标准的方法可证明 (2) 成立. □

由定理 2.3.3, 存在以 $\mathbf{R}^d$ 为状态空间的正规 Markov 过程 X 以热核为转移半群. 下面我们将证明存在 X 的一个修正, 其几乎所有轨道是连续的, 那么这个修正称为是 d-维 Brown 运动. 精确的定义如下.

定义 2.6.1 以 $\mathbf{R}^d$ 为状态空间的时齐 Markov 过程

$$
B = (\Omega, \mathscr{F}, \mathscr{F}_t, B_t, \theta_t, \mathbb{P}^x)
$$

称为是 d-维 $(\mathscr{F}_t)$-标准 Brown 运动, 如果:

(1) 存在 $\Omega_0 \in \mathscr{F}$ 满足对任何 $x \in \mathbf{R}^d$, $\mathbb{P}^x(\Omega_0) = 1$, 使得对 $\omega \in \Omega_0$, $t \mapsto B_t(\omega)$ 是 $[0, +\infty)$ 到 $\mathbf{R}^d$ 的连续函数;

(2) 对任何 $A \in \mathscr{B}(\mathbf{R}^d)$ 且 $t > s \geqslant 0$, 有

$$
\mathbb{P}^x(B_t - B_s \in A | \mathscr{F}_s) = \int_A \frac{\mathrm{e}^{-\frac{|y|^2}{2(t-s)}}}{(2\pi(t-s))^{\frac{d}{2}}} \mathrm{d}y.
$$

由例 2.3.2 知道 B 一定是平稳独立增量过程. 从 0 点出发(或者在概率 $\mathbb{P}^0$ 下)的 Brown 运动称为是标准 Brown 运动. 下面我们要证明 Brown 运动的存在性. 实际上是证明存在概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 和随机过程 $B = (B_t)$ 满足:

(1) $B_0 = 0$ a.s.;

(2) B 是连续过程;

(3) B 是独立增量的;

(4) 对任何 $t > s \geqslant 0$, $B_t - B_s \sim N(0, t-s)$.如果这样的过程存在,那么通过引理 2.2.2 所用的轨道映射, 对任何 $x \in \mathbf{R}^d$, 过程 $(x + B_t)$ 在典则空间上的像测度记为 $\mathbb{P}^x$. 容易验证典则过程就是满足定义 2.6.1 的关于自然流的标准 Brown 运动. 因此也简单地把满足这4个条件的过程称为 d-维标准 Brown 运动(参考定理 2.6.1 后面的注2).

设 N 是任何给定的自然数, $\mathbf{D}$ 是 $[0, \infty)$ 上二分点全体, 即

$$\mathbf{D} := \left\{ \frac{k}{2^m} : m, k = 0, 1, 2, \cdots \right\},$$

令 $\mathbf{D}_N := \mathbf{D} \cap [0, N]$. 形如 $\dfrac{k}{2^m}$ 的二分点称为 m 阶二分点.

引理 2.6.2 设 f 是 $\mathbf{D}$ 上的函数, 如果存在常数 $a > 0$, 对任何自然数 N, 存在常数 $C > 0$ 及自然数 u, 使得当 $m \geqslant u$ 时,

$$\left| f\left(\frac{k+1}{2^m}\right) - f\left(\frac{k}{2^m}\right) \right| \leqslant \frac{C}{2^{ma}}, \ k = 0, 1, \cdots, 2^m N - 1,$$

则 f 可以唯一地延拓成 $[0, \infty)$ 上的连续函数, 且是 a 阶局部 Hölder 连续的, 即在任何 $[0, \infty)$ 的有界区间上是 a 阶 Hölder 连续的.

证明 固定自然数 N 及对应的常数 C 与 u, 取任何 $m \geqslant u$, $s, t \in \mathbf{D}_N$ 及 $|s-t| \leqslant \dfrac{1}{2^m}$, 则 s 落在某个区间 $\left[\dfrac{k}{2^m}, \dfrac{k+1}{2^m}\right)$ 中, 我们来计算 $f(s)$ 与 f 在左端点的值 $f\left(\dfrac{k}{2^m}\right)$ 之间的差, 因 $s \in \mathbf{D}_N$, 写出其二进表示, 形式为

$$s = \frac{k}{2^m} + \frac{b_1}{2^{m+1}} + \cdots + \frac{b_l}{2^{m+l}},$$

其中 b_j 取值为 0 或 1, 令

$$s_0 := \frac{k}{2^m}, \ s_j := s_0 + \sum_{i=1}^{j} \frac{b_i}{2^{m+i}}, \ j = 1, 2, \cdots, l.$$

显然 $\dfrac{k}{2^m} = s_0 \leqslant s_1 \leqslant \cdots \leqslant s_l = s$, 且 s_j 与 s_{j-1} 或相同或是 $m+j$ 阶二分点的相邻点, 由定理条件

$$|f(s_j) - f(s_{j-1})| \leqslant \frac{C}{2^{(m+j)a}},\ j = 1, 2, \cdots, l.$$

因此

$$\left|f(s) - f\left(\frac{k}{2^m}\right)\right| \leqslant \sum_{j=1}^{l} |f(s_j) - f(s_{j-1})| \leqslant \sum_{j=1}^{\infty} \frac{C}{2^{(m+j)a}} = \frac{C}{2^{ma}(2^a - 1)},$$

而 t 与 s 的距离不超过 $\frac{1}{2^m}$, 故 t 落在区间

$$\left[\frac{k-1}{2^m}, \frac{k}{2^m}\right),\ \left[\frac{k}{2^m}, \frac{k+1}{2^m}\right),\ \left[\frac{k+1}{2^m}, \frac{k+2}{2^m}\right)$$

之一, 无论哪一种情况, $f(t)$ 与 f 在左端点的值之间的差同样不大于 $\frac{C}{2^{ma}(2^a - 1)}$, 因此

$$|f(s) - f(t)| \leqslant \frac{C}{2^{ma}} + \frac{2C}{2^{ma}(2^a - 1)} = \frac{C_1}{2^{ma}},$$

其中常数 $C_1 = C + \frac{2C}{2^a - 1}$. 这样, 我们证明了对任何 $m \geqslant u$, $s, t \in \mathbf{D}_N$ 且 $|s - t| \leqslant \frac{1}{2^m}$, 有 $|f(s) - f(t)| \leqslant \frac{C_1}{2^{ma}}$, 从而对 $s, t \in \mathbf{D}_N$, 当 $|s - t| \leqslant \frac{1}{2^u}$ 时, $|f(s) - f(t)| \leqslant C_1|s - t|^a$, 其中常数 u, C 与 N 有关.

现在对任何 $x \in [0, N]$, 必有点列 $\{s_n\} \subset \mathbf{D}_N$, 使得 $s_n \to x$, 则 $\{f(s_n)\}$ 是 Cauchy 列, 记其极限为 $F_N(x)$, 此极限值与收敛点列 $\{s_n\}$ 的选取无关, 因此 F_N 是 $[0, N]$ 上的函数, 是 f 限制在 $\mathbf{D}_N$ 上的连续延拓, 显然这个延拓是唯一的. 对任何 $x, y \in [0, N]$ 且 $|x - y| \leqslant \frac{1}{2^{u+1}} < \frac{1}{2^u}$, 取点列 $\{s_n\}, \{t_n\} \subset \mathbf{D}_N$ 且 $s_n \to x$, $t_n \to y$, 则对充分大的 n, $|s_n - t_n| \leqslant \frac{1}{2^u}$, 故 $|f(s_n) - f(t_n)| \leqslant C_1|s_n - t_n|^a$, 取极限得

$$|F_N(x) - F_N(y)| \leqslant C_1|x - y|^a,$$

即 F_N 是 $[0, N]$ 上 a 阶 Hölder 连续的.

显然 $\{F_N : N \geqslant 1\}$ 是相容的, 即 F_{N+1} 限制在 $[0, N]$ 上与 F_N 是一致的, 对任何 $x \in [0, \infty)$, 如 $x < N$, 定义 $F(x) := F_N(x)$, 由相容性, 定义无歧义, 上面的讨论

证明了 F 是 f 的唯一延拓, 且 F 是局部 a 阶 Hölder 连续的. □

定理 2.6.1(Wiener) 对任何 $d \geqslant 1$, d-维 Brown 运动是存在的.

证明 首先设概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上随机过程 $X = (X_t : t \geqslant 0)$ 是 $\mathbf{R}^d$ 上热核半群的一个实现. 不丢失一般性, 设 $d = 1$. 先让我们计算增量 $X_t - X_s$, $t > s \geqslant 0$ 的矩, 因热核是对称的, 奇数阶矩都是零, 对整数 $k \geqslant 1$,

$$\mathbb{E}^x |X_t - X_s|^{2k} = \int_{\mathbf{R}} \frac{x^{2k}}{\sqrt{2\pi(t-s)}} \mathrm{e}^{-\frac{x^2}{2(t-s)}} \mathrm{d}x = |t-s|^k C(k),$$

而 $C(k) := \mathbb{E}|X_1 - X_0|^{2k} = (2k-1)!!$, 特别地 $\mathbb{E}^x|X_t - X_s|^4 = 3|t-s|^2$.

取 $a > 0$, 由 Chebyshev 不等式,

$$\mathbb{P}^x\left(|X_{\frac{k+1}{2^m}} - X_{\frac{k}{2^m}}| \geqslant \frac{1}{2^{ma}}\right) \leqslant 2^{4ma}\mathbb{E}|X_{\frac{k+1}{2^m}} - X_{\frac{k}{2^m}}|^4 = \frac{3}{2^{2m-4ma}},$$

然后, 对任何自然数 N,

$$\mathbb{P}^x\left(\bigcup_{k=0}^{2^m N-1}\left\{|X_{\frac{k+1}{2^m}} - X_{\frac{k}{2^m}}| \geqslant \frac{1}{2^{ma}}\right\}\right) \leqslant \sum_{k=0}^{2^m N-1} \frac{1}{2^{2m-4ma}} = \frac{3N}{2^{m-4ma}}.$$

让 $4a < 1$, 则

$$\sum_{m\geqslant 0} \mathbb{P}^x\left(\bigcup_{k=0}^{2^m N-1}\left\{|X_{\frac{k+1}{2^m}} - X_{\frac{k}{2^m}}| \geqslant \frac{1}{2^{ma}}\right\}\right) < \infty.$$

由 Borel-Cantelli 引理, 若记

$$\Omega_0^N := \liminf_m \bigcap_{k=0}^{2^m N-1}\left\{|X_{\frac{k+1}{2^m}} - X_{\frac{k}{2^m}}| \leqslant \frac{1}{2^{ma}}\right\},$$

则 $\Omega_0^N \in \mathscr{F}$ 且 $\mathbb{P}^x(\Omega_0^N) = 1$; 再记 $\Omega_0 := \bigcap_N \Omega_0^N$, 则 $\mathbb{P}^x(\Omega_0) = 1$.

对任何 $\omega \in \Omega_0$, 轨道函数 $t \mapsto X_t(\omega)$ 满足: 对任何自然数 N, 存在 $u = u(\omega, N)$, 使得当 $m \geqslant u$ 时,

$$|X_{\frac{k+1}{2^m}}(\omega) - X_{\frac{k}{2^m}}(\omega)| \leqslant \frac{1}{2^{ma}},\ k = 0, 1, 2, \cdots, 2^m N - 1.$$

由引理 2.6.2, $X_t(\omega), t \in \mathbf{D}$ 有唯一连续延拓, 记为 $B_t(\omega), t \in [0, \infty)$, 它在 $[0, \infty)$

上连续. 对 $\omega \notin \Omega_0$, 定义 $B_t(\omega) = 0$, 显然对固定的 $t \geqslant 0$, 取点列 $\{t_n\} \subset \mathbf{D}$ 且 $t_n \to t$, 有

$$B_t = \lim_{n\to\infty} X_{t_n} \cdot 1_{\Omega_0}.$$

因此 B_t 是随机变量, 故 $B = (B_t : t \geqslant 0)$ 是一个实值随机过程. 另外, B_t 是 $\{X_{t_n}\}$ 的几乎处处收敛的极限, 而 $\mathbb{E}^x|X_{t_n} - X_t|^2 = |t_n - t|$, 即 X_t 是 $\{X_{t_n}\}$ 的 L^2 收敛的极限, 故 $B_t = X_t$ a.s.. □

注 1　定理中阶数 a 来自于等式 $\mathbb{E}^x|X_t - X_s|^4 = 3|t-s|^2$, 我们可使用等式 $\mathbb{E}^x|X_t - X_s|^{2k} = C(k)|t-s|^k$, 得到 $a < \dfrac{k-1}{2k}$, 因此 B 的轨道可以是任意 $a < \dfrac{1}{2}$ 阶局部 Hölder 连续的.

注 2　设 Ω 是 $\mathbf{R}_+$ 到 $\mathbf{R}^d$ 的连续映射全体, (B_t) 是坐标过程, $\mathscr{F}_t = \sigma(\{B_s : s \leqslant t\})$, $\mathscr{F} = \mathscr{F}_\infty$. 因为 Brown 运动是平稳独立增量的, 它等价于空间齐性, 故我们经常只需考虑 0 点出发的轨道, 即证明 $(\Omega, \mathscr{F})$ 存在概率测度 $\mathbb{P}$, 使得 $B = (B_t)$ 是标准 Brown 运动. 对任何 $x \in \mathbf{R}^d$, $\omega \in \Omega$, 定义 $\xi_x\omega(t) = \omega(t) + x$. 显然 ξ_x 是 Ω 上可测映射, 记 $\mathbb{P}^x := \mathbb{P}\circ\xi_x$, 那么在 $\mathbb{P}^x$ 下, $(B_{t_1}, B_{t_2}, \cdots, B_{t_n})$, $0 \leqslant t_1 < t_2 < \cdots < t_n$, 的联合分布为

$$\mathbb{E}^x(f(B_{t_1}, \cdots, B_{t_n})) = \int_{x_1\in\mathbf{R}^d, \cdots, x_n\in\mathbf{R}^d} f(x_1, x_2, \cdots, x_n)$$

$$p(t_1, x_1 - x)p(t_2 - t_1, x_2 - x_1)\cdots p(t_n - t_{n-1}, x_n - x_{n-1})\,\mathrm{d}x_1\cdots\mathrm{d}x_n,$$

其中 f 是 $\mathbf{R}^d \times \cdots \times \mathbf{R}^d$ 上的有界 Borel 可测函数. 另外对 $t \geqslant 0$, 定义推移算子 $\theta_t\omega(s) := \omega(t+s)$, $s \geqslant 0$. 容易验证 θ_t 也是 Ω 上的可测映射, 由此构造的组

$$B = (\Omega, \mathscr{F}, (\mathbb{P}^x), (B_t), (\theta_t))$$

自然是定义 2.6.1 意义下的 Brown 运动, 我们称其是 $\mathbf{R}^d$ 上典则 Brown 运动.

注 3　考虑空间 $W := C([0,1], \mathbf{R}^d)$, $[0,1]$ 到 $\mathbf{R}^d$ 的连续映射全体, $\mathscr{B}(W)$ 是 W 上由柱集生成的 σ- 代数, 那么 Brown 运动的存在性定理也证明了 $(W, \mathscr{B}(W))$ 上存在一个概率测度 μ, 满足对任何 $0 < t_1 < \cdots < t_n \leqslant 1$, $A_1, \cdots, A_n \in \mathscr{B}(W)$, 有

$$\mu(\{x \in W : x(t_1) \in A_1, \cdots, x(t_n) \in A_n\}) = \mathbb{P}(B_{t_1} \in A_1, \cdots, B_{t_n} \in A_n).$$

测度 μ 称为是 Wiener 测度, 空间 $(W,\mathscr{B}(W),\mu)$ 称为 Wiener 空间.

上面我们证明了 Brown 运动作为一个有连续轨道和热核半群的随机过程的存在性. Brown 运动有许多好的性质. 首先, 容易看出标准 Brown 运动是一个中心化的 Gauss 过程(见§2.1), 且 $\mathbb{E}(B_tB_s)=t\wedge s$. 反之, 因为中心化 Gauss 过程的有限维分布由其协方差矩阵唯一决定, 故我们有下列刻画.

定理 2.6.2 一个连续实值随机过程 B 是标准 Brown 运动当且仅当它是中心化的 Gauss 过程且 $\mathbb{E}(B_tB_s)=t\wedge s$.

由此容易验证下列关于 Brown 运动的性质, 详细的证明留给读者.

推论 2.6.1 设 $B=(B_t:t\geqslant 0)$ 是一个 Brown 运动.

(1) Markov 性: 对 $s\geqslant 0$, $(B_{t+s}-B_s:\ t\geqslant 0)$ 是一个独立于 $\mathscr{F}_s^0$ 的标准 Brown 运动;

(2) 自相似性: 对任何 $c\neq 0$, $(\frac{1}{c}B_{c^2t}:\ t\geqslant 0)$ 是一个 Brown 运动;

(3) 0 与 ∞ 的对称性: 设 B 是标准的, 定义

$$B'_t=\begin{cases} tB_{\frac{1}{t}}, & t>0,\\ 0, & t=0.\end{cases}$$

则 $B'=(B'_t:t\geqslant 0)$ 也是标准 Brown 运动.

虽然 Brown 运动的轨道是连续的, 但下面的定理说明它们是处处不可导的.

定理 2.6.3 设 B 是定义在概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 上的 Brown 运动, 则存在 $\Omega_0\in\mathscr{F}$, $\mathbb{P}(\Omega_0)=1$, 使得对任何 $\omega\in\Omega_0$, $t\mapsto B_t(\omega)$ 在 $[0,\infty)$ 的任何有界区间上都不是有界变差的.

证明 设 $t>s\geqslant 0$,

$$\begin{aligned}&\mathbb{E}[(B_t-B_s)^2-(t-s)]^2\\ =\ &\mathbb{E}[(B_t-B_s)^4-2(t-s)(B_t-B_s)^2+(t-s)^2]\\ =\ &2(t-s)^2.\end{aligned}$$

另外, 设 $t_2>s_2\geqslant t_1>s_1\geqslant 0$, 则由独立增量性,

$$\mathbb{E}[(B_{t_2}-B_{s_2})^2-(t_2-s_2)][(B_{t_1}-B_{s_1})^2-(t_1-s_1)]=0.$$

现在, 固定 $t>s\geqslant 0$,

$$\Delta_n:=\{s=t_{n,0}<t_{n,1}<\cdots<t_{n,k_n}=t\}$$

是 $[s,t]$ 的一个划分, 且当 $n\to\infty$ 时,

$$|\Delta_n|:=\max|t_{n,j}-t_{n,j-1}|\to 0.$$

记 $\Delta B_{t_{n,j}}:=B_{t_{n,j}}-B_{t_{n,j-1}}$, $\Delta t_{n,j}=t_{n,j}-t_{n,j-1}$. 计算得

$$\begin{aligned}\mathbb{E}\left(\sum_{j=1}^{k_n}(\Delta B_{t_{n,j}})^2-(t-s)\right)^2&=\mathbb{E}\left(\sum_{j=1}^{k_n}[(\Delta B_{t_{n,j}})^2-\Delta t_{n,j}]\right)^2\\&=\sum_{j=1}^{k_n}\mathbb{E}\left((\Delta B_{t_{n,j}})^2-\Delta t_{n,j}\right)^2\\&=\sum_{j=1}^{k_n}2(\Delta t_{n,i})^2\leqslant 2(t-s)|\Delta_n|\to 0,\end{aligned}$$

即 $\sum\limits_{j=1}^{k_n}(B_{t_{n,j}}-B_{t_{n,j-1}})^2$ 平方收敛于 $t-s$. 因此存在划分列 $\{\Delta_n\}$ 的一个子列, 仍然记为 $\{\Delta_n\}$, 使得

$$\sum_{j=1}^{k_n}(B_{t_{n,j}}-B_{t_{n,j-1}})^2\to t-s\quad\text{a.s.},$$

则因为 B 连续, 使上式收敛的轨道在 $[s,t]$ 一定不是有界变差的. 现在将不收敛的 ω 组成的集合记为 $N_{s,t}$, 容易验证如果 $\omega\notin N_{s,t}$, $r\mapsto B_r(\omega)$ 在 $[s,t]$ 上不是有界变差的. 记

$$N:=\bigcup_{s,t\in\mathbf{Q},0\leqslant s<t}N_{s,t},\ \Omega_0:=\Omega\backslash N,$$

其中 $\mathbf{Q}$ 是有理数集, 则 $\mathbb{P}(\Omega_0)=1$, 且对任何 $\omega\in\Omega_0$, $t\mapsto B_t(\omega)$ 在任何的有界区间上都不是有界变差的. □

习 题

1. 设 $r<s<t$. 求 $\mathbb{E}^0(B_s|B_r,B_t)$.

2. (平移性质)设 $B=(\Omega,\mathscr{F},\{\mathbb{P}^x:x\in\mathbf{R}^d\},(B_t)_{t\in\mathsf{T}})$ 是 d-维 Brown 运动. 对 $x\in\mathbf{R}^d$, 定义 $\gamma_x:\Omega\to\Omega$ 满足 $B_t(\gamma_x\omega)=B_t(\omega)+x,\ t\in\mathsf{T},\omega\in\Omega$.

   (1) 证明: 这样的映射存在. 称为是 Ω 上的平移算子;

   (2) 证明: $\mathbb{P}\circ\gamma_x^{-1}=\mathbb{P}^x$.

3. (旋转性质)设 $B=(\Omega,\mathscr{F},\{\mathbb{P}^x:x\in\mathbf{R}^d\},(B_t)_{t\in\mathsf{T}})$ 是典则空间上 d-维 Brown 运动, O 是 $\mathbf{R}^d$ 上的正交变换. 定义 Ω 到自身的映射, 仍然用 O 表示:

$$(O\omega)(t):=O(\omega(t)),\ t\in\mathsf{T},\omega\in\Omega.$$

   证明: O 是可测的且 $\mathbb{P}^x\circ O^{-1}=\mathbb{P}^{Ox}$, $x\in\mathbf{R}^d$, 因此 $\mathbb{P}^0$ 是正交变换下不变的.

4. 对 $r>0$, 令 $T_r:=\inf\{t:|B_t-B_0|\geqslant r\}$. 对 $x\in\mathbf{R}^d$, 证明:

   (1) $\mathbb{P}^x(T_r<\infty)=1$;

   (2) 分布 $\mathbb{P}^x(B_{T_r}\in\cdot)$ 是球面 $\{y:|y-x|=r\}$ 上均匀分布.

5. 证明定理 2.6.2 和定理2.6.1.

6. 一个 $\mathscr{F}$ 的子 σ- 代数称为有 0-1 律, 如果它仅含有概率为 0 或 1 的集合. 设 $B=(\Omega,\mathscr{F},\{\mathbb{P}^x:x\in\mathbf{R}^d\},(B_t)_{t\in\mathsf{T}})$ 是 d-维 Brown 运动. 证明: $\bigcap_{t>0}\sigma(\{B_s:s\geqslant t\})$ 有 0-1 律. 问 $\bigcap_{t>0}\sigma(\{B_s:s\leqslant t\})$ 是否有 0-1 律?

7. 设 $B=(B_t)_{t\in\mathsf{T}}$ 是 1-维 Brown 运动. 证明:

$$\mathbb{P}(\sup_{t>0}B_t=+\infty,\ \inf_{t>0}B_t=-\infty)=1.$$

8. 设 $B=(B_t)_{t\in\mathsf{T}}$ 是 1-维 Brown 运动. 证明: 存在 $t_n\downarrow 0$, 使得

$$\mathbb{P}(\bigcap_n\bigcup_{k\geqslant n}\{B_{t_k}>0\},\ \bigcap_n\bigcup_{k\geqslant n}\{B_{t_k}<0\})=1.$$

9. 证明: 对 $x>0$ 有

$$\int_x^\infty \mathrm{e}^{-\frac{z^2}{2}}\mathrm{d}z\leqslant\frac{1}{x}\mathrm{e}^{-\frac{x^2}{2}},$$

   且当 $x\to\infty$ 时, 两者是等价无穷小.

10. (重对数律) 设 $B=(B_t)_{t\in\mathsf{T}}$ 是 1-维 Brown 运动. 证明:

$$\mathbb{P}\left[\limsup_{t\to 0}\frac{B_t}{\sqrt{2t\ln\ln\dfrac{1}{t}}}=1\right]=1;$$

$$\mathbb{P}\left[\limsup_{t\to \infty}\frac{B_t}{\sqrt{2t\ln\ln t}}=1\right]=1.$$

11. 设 B 是 d-维标准 Brown 运动, 证明: 对任何 $x\in\mathbf{R}^d$, $\|x\|=1$, $\langle x,X_t\rangle$ 是 1-维标准 Brown 运动.

12. 设 $B=(B(t):t\in\mathsf{T})$ 是 1-维标准 Brown 运动, 令

$$X_t:=\mathrm{e}^{-t}B(\mathrm{e}^{2t}),\ t\in\mathbf{R}.$$

证明: X 是具有 Markov 性的平稳的中心化 Gauss 过程, 并求出其协方差函数. 过程 X 称为是 Ornstein-Ulenbeck 过程.

13. 设 $\mathbb{P}$ 是 $(\mathbf{R}^{\mathsf{T}},\mathscr{B}^{\mathsf{T}})$ 上概率测度, 其有限维分布与 Brown 运动的一致. C 是 $\mathbf{R}^{\mathsf{T}}$ 中的连续映射全体. 证明: $\mathbb{P}_*(C)=0$, $\mathbb{P}^*(C)=1$, 其中 $\mathbb{P}_*$ 与 $\mathbb{P}^*$ 分别是 $\mathbb{P}$ 引导的内测度与外测度.

14. Brown 运动还可以用 Fourier 级数的方法产生. 设 $\{\xi_n:n\geqslant 0\}$ 是概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 上的都服从标准正态分布的独立随机变量序列. (说明其存在性.) 对任何 $f\in L^2([0,\pi])$, 令

$$H(f):=a_0\xi_0+a_1\xi_1+a_2\xi_2+\cdots,$$

其中 $\{a_n\}$ 是 f 的 Fourier 系数: $a_n:=\dfrac{1}{\pi}\int_0^\pi f(x)\cos nx\mathrm{d}x$. 显然

$$\mathbb{E}H(f)^2=\sum_{n\geqslant 0}a_n^2=\int f^2(x)\mathrm{d}x,$$

即 H 是 $L^2([0,\pi])$ 到 $L^2(\Omega,\mathscr{F},\mathbb{P})$ 的一个等距嵌入. 对 $t\in[0,\pi]$, 令

$$X_t:=H(1_{[0,t]})=\frac{t}{\sqrt{\pi}}\xi_0+\sum_{n\geqslant 1}\frac{\sin nt}{n}\xi_n,$$

那么

(1) 以上级数 L^2 收敛且 a.s. 收敛;

(2) $(X_t : t \in [0,\pi])$ 是中心化 Gauss 过程且 $\mathbb{E}X_tX_s = t \wedge s$;

(3) 以上级数在 $t \in [0,\pi]$ 以概率 1 一致收敛. 因此 (X_t) 是连续过程, 推出它在 $[0,\pi]$ 上是 Brown 运动.

第三章 随机分析基础

鞅起源于赌博游戏, 它是指一个无偏向的赌博规则. 现在鞅是现代随机分析中的重要工具之一, 其系统的并让概率学家们看到其重要性的研究要归功于 J.L. Doob 在 20 世纪中叶的工作. 本章的前几节将着重讨论鞅的正则化定理, Doob 的鞅不等式及鞅的基本性质. 在这一章的其后几节中, 我们将介绍 Itô 的关于连续鞅的随机积分理论. 前面在 §2.6 中, 我们证明了 Brown 运动的几乎所有轨道在任何区间上都不是有界变差的, 因此不可能按轨道以 Stieltjes 积分的方式来定义关于 Brown 运动的积分, 即通常的以测度为起点的积分理论对 Brown 运动是不适合的, 而随机积分实际上只是两个线性度量空间之间的具有类似于通常积分性质的保距线性映射, 我们还将证明随机积分理论中极其重要的 Itô 公式并且演示它的一些重要应用.

§3.1 σ- 代数流与停时

如果说前面我们所讨论的概率论与随机过程部分都还没有完全脱离测度论的框架, 则下面将引入的停时的概念已完全超越了测度论的范畴. 停时是经典随机过程理论中最能体现概率直观背景的概念之一. 可以说, 没有停时的讨论, 就不能称为随机分析. 设有概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 和 $\mathsf{T} \subset \mathbf{R}$. 回忆 $\mathscr{F}$ 的子 σ- 代数族 $(\mathscr{F}_t : t \in \mathsf{T})$ 称为流, 如果对任何 $s < t$, 有 $\mathscr{F}_s \subset \mathscr{F}_t$. 对任何 $t \in \mathsf{T}$, 定义 $\mathscr{F}_{t+} := \bigcap_{s>t} \mathscr{F}_s$, 那么 $(\mathscr{F}_{t+})$ 也是流.

定义 3.1.1 给定概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 及流 $(\mathscr{F}_t)$, 映射 $\tau : \Omega \longrightarrow \mathsf{T} \cup \{\infty\}$ 称为是一个 $(\mathscr{F}_t)$ 停时, 如果对任何 $t \in \mathsf{T}$, $\{\omega \in \Omega : \tau(\omega) \leqslant t\} \in \mathscr{F}_t$.

停时是相对于一个流而言的, 但当上下文明确时, 就简单地说一个停时. $(\mathscr{F}_t)$ 停时必是 $(\mathscr{F}_{t+})$ 停时. 设 τ 是 $(\mathscr{F}_t)$ 停时, 容易验证 $\mathscr{F}$ 中使得 $A \cap \{\tau \leqslant t\} \in \mathscr{F}_t$ 对所有 $t \in \mathsf{T}$ 成立的 A 全体是一个 σ 代数, 记为 $\mathscr{F}_\tau$. 显然如果 $\tau \equiv t$ 时, $\mathscr{F}_\tau = \mathscr{F}_t$,

另外一个随机变量 ξ 是 $\mathscr{F}_\tau$ 可测当且仅当对任何 $t \in \mathsf{T}$, $\xi \cdot 1_{\{\tau \leqslant t\}}$ 是 $\mathscr{F}_t$ 可测的.

引理 3.1.1 设 τ, σ, τ_n 是 $(\mathscr{F}_t)$ 停时.

(1) $\tau \vee \sigma$, $\tau \wedge \sigma$ 是停时;

(2) 则 τ 是 $(\mathscr{F}_{t+})$ 停时当且仅当对任何 $t \in \mathsf{T}$, $\{\tau < t\} \in \mathscr{F}_t$;

(3) 当 τ_n 单调上升(或下降, 且 $(\mathscr{F}_t)$ 右连续)时, $\lim \tau_n$ 是停时;

(4) 如果 $\sigma \leqslant \tau$, 则 $\mathscr{F}_\sigma \subset \mathscr{F}_\tau$;

(5) 如果 $(\mathscr{F}_t)$ 右连续且 $\tau_n \downarrow \tau$, 则 $\bigcap_n \mathscr{F}_{\tau_n} = \mathscr{F}_\tau$;

(6) τ 是 $\mathscr{F}_\tau$ 可测的.

证明 (1) 显然 $\{\tau \vee \sigma \leqslant t\} = \{\tau \leqslant t\} \cap \{\sigma \leqslant t\}$, 而 $\{\tau \wedge \sigma \leqslant t\} = \{\tau \leqslant t\} \cup \{\sigma \leqslant t\}$.

(2) 设 τ 是 $(\mathscr{F}_{t+})$ 停时, 则 $\{\tau < t\} = \bigcup_n \{\tau \leqslant t - \dfrac{1}{n}\} \in \mathscr{F}_t$. 反之, 对任何 k, $\{\tau \leqslant t\} = \bigcap_{n>k} \{\tau < t + \dfrac{1}{n}\} \in \mathscr{F}_{t+\frac{1}{k}}$, 故 $\{\tau \leqslant t\} \in \mathscr{F}_{t+}$, 即 τ 是 $(\mathscr{F}_{t+})$ 停时.

(3) 显然若 τ_n 单调上升, $\{\lim \tau_n \leqslant t\} = \bigcup_n \{\tau_n \leqslant t\}$; 反之若 τ_n 单调下降且 $(\mathscr{F}_t)$ 右连续, 由 (2), $\{\lim \tau_n < t\} = \bigcap_n \{\tau_n < t\} \in \mathscr{F}_t$, 故 $\lim \tau_n$ 是停时.

(4) 设 $A \in \mathscr{F}_\sigma$, 对 $t \in \mathsf{T}$, 有 $A \cap \{\tau \leqslant t\} = (A \cap \{\sigma \leqslant t\}) \cap \{\tau \leqslant t\} \in \mathscr{F}_t$, 故 $A \in \mathscr{F}_\tau$.

(5) 显然 $\bigcap_n \mathscr{F}_{\tau_n} \supset \mathscr{F}_\tau$, 反之设 $A \in \bigcap_n \mathscr{F}_{\tau_n}$, 则对 $t \in \mathsf{T}$,

$$A \cap \{\tau < t\} = \bigcup_n (A \cap \{\tau_n < t\}) \in \mathscr{F}_t.$$

即 $A \in \mathscr{F}_\tau$.

(6) 对任何 $t, s \in \mathsf{T}$, $\{\tau \leqslant t\} \cap \{\tau \leqslant s\} = \{\tau \leqslant s \wedge t\} \in \mathscr{F}_{s \wedge t} \subset \mathscr{F}_t$, 因此对任何 $s \in \mathsf{T}$, $\{\tau \leqslant s\} \in \mathscr{F}_\tau$. □

停时是一个在日常生活中大家普遍使用的一个概率论的概念之一. 如在一个赌博中称在输或赢到一定数量的钱后离开, 这与说赌若干次后停止有本质的区别, 因为前者是一个随机时间. 停时是静止的时间 t 的推广, 它的原型是如下所言的进入时与首中时.

例 3.1.1 首先考虑一个简单的例子. 设 X 是一个 Markov 链, 对 $x \in \mathbf{Z}$, 如 §2.4 中所定义的, τ_x 是 X 首次到达 x 的时间, 即

$$\tau_x := \inf\{n \geqslant 1:\ X_n = x\}.$$

因为 $\{\tau_x \leqslant n\} = \bigcup_{k=1}^n \{X_k = x\}$, 故 τ_x 是一个停时. 一般地, 对 $I \subset \mathbf{Z}$, 令 $\tau_I := \inf\{n \geqslant 1:\ X_n \in I\}$. 显然 τ_I 是停时且 $\tau_I = \inf_{x \in I} \tau_x$.

例 3.1.2 设 E 是一个度量空间, X 是关于流 $(\mathscr{F}_t)$ 适应的以 E 为状态空间的随机过程. 对 $A \subset E$, $\omega \in \Omega$, 定义

$$D_A(\omega) := \inf\{t \geqslant 0 : X_t(\omega) \in A\};$$

$$T_A(\omega) := \inf\{t > 0 : X_t(\omega) \in A\}.$$

我们先证明若 A 是闭集且 X 是连续过程, 则 D_A 是 $(\mathscr{F}_t)$ 停时. 事实上, 这时由连续性 $D_A \leqslant t$ 当且仅当

$$\inf_{s \in [0,t] \cap \mathbf{Q}} d(X_s, A) = d(\{X_s : s \in [0,t] \cap \mathbf{Q}\}, A) = 0.$$

因此 $\{D_A \leqslant t\} \in \mathscr{F}_t$. 另一方面, 若 A 是开集且 X 是右连续的, 由右连续性, $T_A < t$ 当且仅当存在有理数 $s < t$, 使得 $X_s \in A$, 因此 $\{T_A < t\} \in \mathscr{F}_t$, 即 T_A 是 $(\mathscr{F}_{t+})$-停时, 但一般不是 $(\mathscr{F}_t)$-停时. 直观地, X 的一条轨道 t 时刻(包括 t)前的信息不能告诉我们它是否将立刻进入一个开集.

随机时间 D_A 与 T_A 分别称为是集合 A 的进入时与首中时, 它们的区别在于过程的初始位置, 若轨道的起始点不在 A 中, 则 $D_A = T_A$; 若轨道从 A 中的点出发, 则 $D_A = 0$ 而 T_A 不一定. 因此, 不管轨道在什么点出发, $T_A < \infty$ 等价于轨道将再回到 A.

例 3.1.3 X 如上, 对 $A \subset E$, $\omega \in \Omega$, 定义

$$L_A(\omega) := \sup\{t > 0 : X_t(\omega) \in A\}.$$

L_A 是轨道最后一次在 A 中的时间, 称为是 A 的末离时. 一般地 L_A 不是停时, 因为轨道在 t 时刻后不再进入 A 这样的事件不能用轨道在 t 时刻前的信息来判断.

定理 3.1.1 设 τ, σ 是 $(\mathscr{F}_t)$ 停时, 则 $\mathscr{F}_{\tau\wedge\sigma} = \mathscr{F}_\tau \cap \mathscr{F}_\sigma$, 且事件 $\{\tau < \sigma\}, \{\tau \leqslant \sigma\} \in \mathscr{F}_\tau \cap \mathscr{F}_\sigma$.

证明 显然由引理 3.1.1 (4), $\mathscr{F}_{\tau\wedge\sigma} \subset \mathscr{F}_\tau \cap \mathscr{F}_\sigma$. 为证另一方向的包含关系, 取 $A \in \mathscr{F}_\tau \cap \mathscr{F}_\sigma$, 对任何 $t \in \mathsf{T}$,

$$
\begin{aligned}
A\cap\{\tau\wedge\sigma\leqslant t\} &= A\cap(\{\tau\leqslant t\}\cup\{\sigma\leqslant t\})\\
&= (A\cap\{\tau\leqslant t\})\cup(A\cap\{\sigma\leqslant t\})\in\mathscr{F}_t.
\end{aligned}
$$

因此 $A\in\mathscr{F}_{\tau\wedge\sigma}$.

对任何 $t\in\mathsf{T}$,

$$
\begin{aligned}
\{\tau<\sigma\}\cap\{\sigma\leqslant t\} &= (\bigcup_{s\in\mathbf{Q}}\{\tau<s\}\cap\{\sigma>s\})\cap\{\sigma\leqslant t\}\\
&= \bigcup_{s\in\mathbf{Q}}(\{\tau<s\}\cap\{\sigma>s\}\cap\{\sigma\leqslant t\})\\
&= \bigcup_{s\in\mathbf{Q},s<t}(\{\tau<s\}\cap\{\sigma>s\}\cap\{\sigma\leqslant t\})\in\mathscr{F}_t.
\end{aligned}
$$

因此 $\{\tau<\sigma\}\in\mathscr{F}_\sigma$. 另外

$$
\begin{aligned}
\{\tau<\sigma\}\cap\{\tau\leqslant t\} &= \{\tau<\sigma,\tau\leqslant t,\sigma\leqslant t\}\cup\{\tau<\sigma,\tau\leqslant t,\sigma>t\}\\
&= \{\tau<\sigma,\sigma\leqslant t\}\cup\{\tau\leqslant t,\sigma>t\}\in\mathscr{F}_t.
\end{aligned}
$$

因此 $\{\tau<\sigma\}\in\mathscr{F}_{\sigma\wedge\tau}$. 对称性推出 $\{\tau>\sigma\}\in\mathscr{F}_{\sigma\wedge\tau}$, 进而 $\{\tau\leqslant\sigma\},\{\sigma\leqslant\tau\}\in\mathscr{F}_{\sigma\wedge\tau}$. □

定理 3.1.2　设 σ,τ 是停时, X 是 $\mathscr{F}_\tau$ 可测的可积随机变量, 则

$$\mathbb{E}(X|\mathscr{F}_\sigma)=\mathbb{E}(X|\mathscr{F}_{\sigma\wedge\tau}).$$

等价地说, 若 X 可积, 则

$$\mathbb{E}[\mathbb{E}(X|\mathscr{F}_\tau)|\mathscr{F}_\sigma]=\mathbb{E}(X|\mathscr{F}_{\sigma\wedge\tau}).$$

证明　以上两个公式的等价性是显而易见的. 我们先证明如 X 是 $\mathscr{F}_\tau$ 可测的, 则 $1_{\{\sigma>\tau\}}X$ 与 $1_{\{\sigma\geqslant\tau\}}X$ 是 $\mathscr{F}_{\sigma\wedge\tau}$ 可测的. 事实上对任何 $t\in\mathsf{T}$,

$$1_{\{\sigma>\tau\}}X\cdot 1_{\{\sigma\wedge\tau\leqslant t\}}=1_{\{\sigma>\tau\}}\cdot X\cdot 1_{\{\tau\leqslant t\}},$$

因 $\{\sigma>\tau\}\in\mathscr{F}_\tau$ 且 X 是 $\mathscr{F}_\tau$ 可测的, 故上式是 $\mathscr{F}_t$ 可测的, 即 $1_{\{\sigma>\tau\}}X$ 是 $\mathscr{F}_{\sigma\wedge\tau}$ 可测的, 同理 $1_{\{\sigma\geqslant\tau\}}X$ 也是 $\mathscr{F}_{\sigma\wedge\tau}$ 可测的.

现在

$$\begin{aligned}\mathbb{E}(X|\mathscr{F}_\sigma) &= \mathbb{E}(1_{\{\sigma>\tau\}}X|\mathscr{F}_\sigma) + \mathbb{E}(1_{\{\tau\geqslant\sigma\}}X|\mathscr{F}_\sigma)\\ &= 1_{\{\sigma>\tau\}}X + 1_{\{\tau\geqslant\sigma\}}\mathbb{E}(X|\mathscr{F}_\sigma),\end{aligned}$$

首先 $1_{\{\sigma>\tau\}}X$ 是 $\mathscr{F}_{\sigma\wedge\tau}$ 可测的, 其次因 $\mathbb{E}(X|\mathscr{F}_\sigma)$ 是 $\mathscr{F}_\sigma$ 可测的, 故 $1_{\{\tau\geqslant\sigma\}}\mathbb{E}(X|\mathscr{F}_\sigma)$ 是 $\mathscr{F}_{\sigma\wedge\tau}$ 可测的. 因此

$$\begin{aligned}\mathbb{E}(X|\mathscr{F}_\sigma) &= \mathbb{E}(1_{\{\sigma>\tau\}}X|\mathscr{F}_{\sigma\wedge\tau}) + \mathbb{E}(1_{\{\tau\geqslant\sigma\}}\mathbb{E}(X|\mathscr{F}_\sigma)|\mathscr{F}_{\sigma\wedge\tau})\\ &= \mathbb{E}(1_{\{\sigma>\tau\}}X|\mathscr{F}_{\sigma\wedge\tau}) + 1_{\{\tau\geqslant\sigma\}}\mathbb{E}(\mathbb{E}(X|\mathscr{F}_\sigma)|\mathscr{F}_{\sigma\wedge\tau})\\ &= 1_{\{\sigma>\tau\}}\mathbb{E}(X|\mathscr{F}_{\sigma\wedge\tau}) + 1_{\{\tau\geqslant\sigma\}}\mathbb{E}(X|\mathscr{F}_{\sigma\wedge\tau})\\ &= \mathbb{E}(X|\mathscr{F}_{\sigma\wedge\tau}).\end{aligned}$$

完成证明. □

要深入研究涉及停时的随机过程的可测性, 我们自然需要引入关于 (t,ω) 可测的概念.

定义 3.1.2　设 $\mathsf{T}=[0,\infty)$, $X=(X_t)_{t\in\mathsf{T}}$ 是以拓扑空间 E 为状态空间的 $(\mathscr{F}_t)_{t\in\mathsf{T}}$ 适应过程, 称过程 X 是可测的, 如果映射 $(s,\omega)\mapsto X_s(\omega)$ 是从 $(\mathsf{T}\times\Omega,\mathscr{B}(\mathsf{T})\times\mathscr{F})$ 到 $(E,\mathscr{E})$ 的可测映射; 称 X 是(关于流$(\mathscr{F}_t)$)循序可测的, 如果对任何 $t\in\mathsf{T}$, 映射 $(s,\omega)\mapsto X_s(\omega)$ 是从 $([0,t]\times\Omega,\mathscr{B}([0,t])\times\mathscr{F}_t)$ 到 $(E,\mathscr{E})$ 的可测映射.

显然, 一个循序可测的过程一定是可测的. 集合 $A\subset\mathsf{T}\times\Omega$ 可以自然地视为一个过程 $(t,\omega)\mapsto 1_A(t,\omega)$, 我们称 A 是循序可测集, 如果它作为过程是循序可测的.

定理 3.1.3　一个右连续(或左连续)的 $(\mathscr{F}_t)$ 适应过程是循序可测的.

证明　对 $t\in\mathsf{T}$, $n\geqslant 1$, 令

$$X_s^{(n)} := \sum_{k=1}^{2^n} X_{\frac{k}{2^n}t}1_{(\frac{k-1}{2^n}t,\frac{k}{2^n}t]} + X_0\cdot 1_{\{0\}},$$

则 $(s,\omega)\mapsto X_s^{(n)}(\omega)$ 是 $([0,t]\times\Omega,\mathscr{B}([0,t])\times\mathscr{F}_t)$ 到 $(E,\mathscr{E})$ 的可测映射. 因 X 是右连续的, 故 $X_s=\lim_n X_s^{(n)}$. 因此 $(s,\omega)\mapsto X_s(\omega)$ 是 $([0,t]\times\Omega,\mathscr{B}([0,t])\times\mathscr{F}_t)$ 到

$(E,\mathscr{E})$ 的可测映射. □

对随机时间 $\tau:\Omega\to\mathsf{T}\cup\{+\infty\}$, 自然地定义

$$X_\tau(\omega):=X_{\tau(\omega)}(\omega),$$

注意, 此映射的定义域是 $\{\tau<\infty\}$, 除非 $X_{+\infty}$ 处处有定义. 当 τ 是 A 的首中时时, X_τ 是首中点. 我们也定义 X 的 τ 停止过程 (或局部化过程)为 $X_t^\tau:=X_{\tau\wedge t}$, $t>0$, 停止过程是我们下面几节内容中所使用的主要技巧之一.

定理 3.1.4 如果 X 是 E-值 $(\mathscr{F}_t)$ 循序可测过程, τ 是 $(\mathscr{F}_t)$ 停时, 则 X_τ 限制在 $\Omega_\tau:=\{\tau<\infty\}$ 上是 $(\Omega_\tau,\mathscr{F}_\tau\cap\Omega_\tau)$ 到 $(E,\mathscr{E})$ 的可测映射.

证明　因 X 是循序可测的, 故对任何 $t\in\mathsf{T}$, $(s,\omega)\mapsto X_s(\omega)$ 是从 $([0,t]\times\Omega,\mathscr{B}([0,t])\times\mathscr{F}_t)$ 到 $(E,\mathscr{E})$ 的可测映射. 而容易验证 $\omega\mapsto(\tau(\omega)\wedge t,\omega)$ 是 $(\Omega,\mathscr{F}_t)$ 到 $([0,t]\times\Omega,\mathscr{B}([0,t])\times\mathscr{F}_t)$ 的可测映射, 因此两个映射的复合 $\omega\mapsto X_{\tau(\omega)\wedge t}(\omega)$ 是 $(\Omega,\mathscr{F}_t)$ 到 $(E,\mathscr{E})$ 的可测映射. 那么对任何 $B\in\mathscr{E}$, $\{X_\tau\in B\}\cap\Omega_\tau\cap\{\tau\leqslant t\}=\{X_\tau\in B,\tau\leqslant t\}=\{X_{\tau\wedge t}\in B\}\cap\{\tau\leqslant t\}\in\mathscr{F}_t$, 即 $\{X_\tau\in B\}\cap\Omega_\tau\in\mathscr{F}_\tau$. □

用 $\mathscr{O}$ (对应地, $\mathscr{P}$) 表示乘积空间 $\mathsf{T}\times\Omega$ 上右连续且具有左极限(对应地, 连续)的适应过程所生成的最小 σ- 代数, 称为是可选(optional) (对应地, 可料(predictable)) σ- 代数, 一个随机过程 X 称为可选的(对应地, 可料的), 如果它是关于 $\mathscr{O}$ (对应地, $\mathscr{P}$) 可测的. 显然可料过程一定是可选的, 上面定理说明可选过程是循序可测的. 上述 σ- 代数都依赖于给定的流 $(\mathscr{F}_t)$.

让我们不加证明地叙述 Meyer 的截面定理, 它是随机分析中最主要的结果.(参考文献 [10].) 在第四章的一些定理的证明中会用到截面定理.

定理 3.1.5 (Section Theorem) (1) 设 Z,Z' 是有界可选随机过程, 如果对任何停时 T 有 $\mathbb{E}(Z_T,T<\infty)=\mathbb{E}(Z'_T,T<\infty)$, 那么 Z 与 Z' 是不可区分的;

(2) 设 Z 是有界可选随机过程, 如果对任何递减有界停时列 $\{T_n\}$ 有

$$\lim_n\mathbb{E}Z_{T_n}=\mathbb{E}Z_{\lim_n T_n},$$

那么 Z 是右连续的.

最后我们要证明一个重要的定理, 即对于一个循序可测过程来说, 任何 Borel 集的进入时和首中时都是停时. 这需要用到测度论中的一个基本定理, 它的证明是

Choquet 容度理论的一个应用, 篇幅较大, 我们在这里仅不加证明地引用, 感兴趣的读者可参考[10]. 首先对任何空间 S, 集合 $A \subset S \times \Omega$ 在 Ω 上的投影定义为

$$\pi(A) := \{\omega \in \Omega : 存在\ s \in S\ 使得\ (s, \omega) \in A\}.$$

设 $(\Omega, \mathscr{F})$ 是可测空间, 对任何概率 μ, $\overline{\mathscr{F}}^{\mu}$ 表示 $\mathscr{F}$ 关于 μ 的完备化, 定义

$$\mathscr{F}^* := \bigcap_{\mu} \overline{\mathscr{F}}^{\mu},$$

其中 μ 取遍所有概率测度. $\mathscr{F}^*$ 称为是 $\mathscr{F}$ 的普遍完备化.

定理 3.1.6 设 $(\Omega, \mathscr{F})$ 是可测空间, $(S, \mathscr{B}(S))$ 是完备可分度量空间及其 Borel σ-代数, 则任何 $A \in \mathscr{B}(S) \times \mathscr{F}$ 的投影 $\pi(A) \in \mathscr{F}^*$.

由此下面的 debut 定理是显然的. 它最初的形式要归功于 Doob 和 Hunt. 说流 $(\mathscr{F}_t)$ 满足通常条件是指它是右连续的且 $\mathscr{F}_0$ 含有 $\mathscr{F}$ 的零概率集及其子集.

定理 3.1.7 设 $(\Omega, \mathscr{F}, \mathbb{P})$ 是完备概率空间, 流 $(\mathscr{F}_t)$ 满足通常条件, $(E, \mathscr{E})$ 是完备可分度量空间及其 Borel σ- 代数.

(1) 若 A 是一个循序集, 则 A 的 debut, $\tau := \inf\{t \geqslant 0 : (t, \omega) \in A\}$, 是 $(\mathscr{F}_t)$ 停时;

(2) 若 X 是一个循序可测过程, $B \in \mathscr{B}(E)$, 则 T_B 是 $(\mathscr{F}_t)$ 停时.

证明 对任何 $t \in \mathsf{T}$, 显然

$$\{\tau < t\} = \pi(([0, t) \times \Omega) \cap A),$$

而 A 是循序的, 故 $([0, t) \times \Omega) \cap A \in \mathscr{B}[0, t] \times \mathscr{F}_t$, 由投影定理且因为 $(\mathscr{F}_t)$ 完备右连续, 推出 τ 是 $(\mathscr{F}_t)$-停时.

最后, 设 X 是一个循序可测过程, $B \in \mathscr{B}(E)$, 令 $A := \{(t, \omega) : X(t, \omega) \in B\}$. 那么 A 是一个循序集且 B 的进入时 D_B 就是 A 的 debut. 因此 Borel 集的进入时是停时. 而

$$T_B = \downarrow \lim_{t \downarrow 0} (t + D_B \circ \theta_t),$$

推出 Borel 集的首中时是停时. $\square$

习题

1. 验证: 若 τ 是停时, 则 $\mathscr{F}_\tau$ 是 σ- 代数且当 $\tau \equiv t$ 时, $\mathscr{F}_\tau = \mathscr{F}_t$.
2. 证明: 随机变量 ξ 是 $\mathscr{F}_\tau$ 可测的当且仅当对任何 $t \in \mathsf{T}, \xi \cdot 1_{\{\tau \leqslant t\}}$ 是 $\mathscr{F}_t$ 可测的.
3. 设 X 是一个 $(\mathscr{F}_t)$ 适应的, 右连续并具有左极限的过程. 令 A 是使得 X 在 $[0, t)$ 上连续的样本全体. 证明: $A \in \mathscr{F}_t$.
4. 设 X 是一个 $(\mathscr{F}_t)$ 适应的, 左连续并具有右极限的过程. 令 A 是使得 X 在 $[0, t]$ 上连续的样本全体. 证明: 如果 $(\mathscr{F}_t)$ 右连续, 则 $A \in \mathscr{F}_t$.
5. 设 X 是随机过程, τ 是停时. 如果对 $\omega, \omega' \in \Omega$, 有 $X_t(\omega) = X_t(\omega')$ 对任何 $t \in [0, \tau(\omega)]$, 证明: $\tau(\omega) = \tau(\omega')$.
6. 设 τ 是个停时, σ 是随机时间且 $\sigma \geqslant \tau$. 证明: 如果 σ 是 $\mathscr{F}_\tau$ 可测的, 则 σ 是个停时.
7. 设 τ 是 $(\mathscr{F}_t)$ 停时, 证明: $\mathscr{F}_{\tau+} = \bigcap_{\sigma > \tau} \mathscr{F}_\sigma$, 其中 σ 是停时.
8. 停时 τ 称为可预料, 如果存在递增停时列 τ_n 使得当 $\tau > 0$ 时几乎处处地 $\tau_n \uparrow\uparrow \tau$. 证明: Brown 运动关于闭集 F 的首中时是可预料的.
9. 停时 τ 称为是绝不可及的, 如果对任何可预料停时 σ 有 $\mathbb{P}(\tau = \sigma) = 0$. 证明: Poisson 过程 N 的首次跳跃

$$\tau := \inf\{t : N_{t-} \neq N_t\}$$

是绝不可及的.

§3.2 鞅与鞅序列

在本节中, 我们将着重介绍鞅的定义及一些常用的例子. 简单地说, 鞅就是公平原则. 在生活中有许多无法预见结果的事件, 如比赛, 掷骰子, 下一个看见的汽车是单号还是双号等. 人们可以在任何这样的事件上进行下注赌博, 只要进行赌博的各方认为规则是公正的. 公正的基本思想是: 风险与可能的获利成正比. 比如买彩票,

中奖的概率极微, 但一旦中奖, 奖额极大, 人们在这里买的是运气, 而不是概率. 又如将钱存入银行, 当然一般不会血本无归, 但一般获利也仅是利息而已. 这就是鞅的基本思想.

设 T 是时间集, $(\mathscr{F}_t)$ 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上给定的 σ- 代数流.

定义 3.2.1 一个可积的 $(\mathscr{F}_t)$ 适应的实值过程称为是关于 $(\mathscr{F}_t)$ 的鞅 (对应地, 下鞅), 如果对任何 $t, s \in \mathsf{T}$, $t > s$, 有 $\mathbb{E}(X_t|\mathscr{F}_s) = X_s$ (对应地, $\mathbb{E}(X_t|\mathscr{F}_s) \geqslant X_s$). 在上下文明确时, 简称为鞅. X 称为是上鞅, 如果 $-X$ 是下鞅. 如果预先未指定一个流, 一个鞅是指关于此过程的自然流的鞅.

鞅与流有关, 关于大的适应流是鞅蕴含着关于小的适应流也是鞅. 因为 X 是上鞅当且仅当 $-X$ 是下鞅, 故我们在此仅需研究鞅与下鞅. 直观地, 对于一个鞅来说, 以到现在为止的信息来预期将来某时刻的输赢是不可能的, 或者说, 至多能知道将来的输赢关于现在的条件期望是零. 当 T 分别是离散时间集或连续时间集时, 对应的鞅分别称为离散时间鞅或连续时间鞅. 显然 $(X_n : n \geqslant 0)$ 是离散时间鞅当且仅当对任何 $n \geqslant 0$, $\mathbb{E}(X_{n+1}|\mathscr{F}_n) = X_n$. 由定义, 立刻得到下列简单性质:

(1) 记 $\mathscr{M}$ (对应地, $\mathscr{M}^+$) 为相对于给定 σ- 代数流的鞅 (对应地, 下鞅) 全体, 显然 $\mathscr{M}$ 是一个线性空间, 而 $\mathscr{M}^+$ 是锥, 且 $\mathscr{M}^+$ 对取大运算 $\vee$ 封闭.

(2) 如果 X 是 $(\mathscr{F}_t)$ 下鞅, 则 $t \mapsto \mathbb{E}X_t$ 递增. 因此, 下鞅 X 是鞅当且仅当 $t \mapsto \mathbb{E}X_t$ 是常数.

(3) 由 Jensen 不等式, 如果 X 是鞅, ϕ 是凸函数, 那么 若 $\phi(X)$ 可积, 则是下鞅. 因此 $|X|$, X^2 (若 X 平方可积) 是下鞅. 另外, 如果 X 是下鞅, ϕ 是下凸递增函数, 那么 $\phi(X)$ 也是下鞅. 因此 X^+ 是下鞅.

例 3.2.1 设 $\{\xi_n : n \geqslant 1\}$ 是一个 Bernoulli 随机序列. 令 $X_0 = 0$, $X_n := \sum\limits_{i=1}^{n} \xi_i$, $n \geqslant 1$, 且 $\{\mathscr{F}_n\}$ 是 X 的自然流, 则对于 $n \geqslant 1$,

$$
\begin{aligned}
\mathbb{E}(X_{n+1}|\mathscr{F}_n) &= \mathbb{E}(X_{n+1} - X_n|\mathscr{F}_n) + X_n \\
&= \mathbb{E}\xi_i + X_n = X_n + (p - q).
\end{aligned}
$$

因此当 $p = q$ 时, X 就是 $\mathbf{Z}$ 上简单随机游动, 是个鞅, $p \geqslant q$ 时, X 是下鞅, $p \leqslant q$ 时, X 是上鞅. 可以看出鞅对应于一个对双方公平的博弈对局, 而下鞅与上鞅分别对

应与一个对己有利与对他有利的博弈对局. (按此结论, 也许把上鞅与下鞅的名称对换一下更适合实际的意义.) ▮

例 3.2.2 设 B 是标准 Brown 运动, $(\mathscr{F}_t)$ 是自然流. 那么对任何 $t > s \geqslant 0$, $B_t - B_s$ 独立于 $\mathscr{F}_s$, 故 $\mathbb{E}(B_t - B_s|\mathscr{F}_s) = 0$, 因此 Brown 运动是鞅. 由同样方法可以验证 $B_t^2 - t$ 和 $\exp(B_t - \frac{1}{2}t)$ 也是鞅. 设有参数为 λ 的 Poisson 过程 N, 则 $N_t - \lambda t$ 是鞅.

一般地, 设 $\mathsf{T} = [0, \infty)$, $X = (X_t)_{t \in \mathsf{T}}$ 是一个可积的独立增量过程, $(\mathscr{F}_t^0)$ 是自然流, 则对于 $t, s \in \mathsf{T}$, $t > s$, 有 $\mathbb{E}(X_t|\mathscr{F}_s^0) = X_s + \mathbb{E}X_t - \mathbb{E}X_s$. 因此当 $t \mapsto \mathbb{E}X_t$ 是单调上升时, X 是下鞅; 单调下降时, X 是上鞅; 常数时, X 是鞅. ▮

例 3.2.3 (Doob 鞅) 设 X 是可积随机变量, 令

$$X_t := \mathbb{E}(X|\mathscr{F}_t),$$

则 $X = (X_t : t \in \mathsf{T})$ 是一个关于 $(\mathscr{F}_t)$ 的鞅. 称这样的鞅 X 为右闭鞅或 Doob 鞅. Doob 鞅是一致可积的, 事实上, 对任何 $N > 0$,

$$\mathbb{E}(|X_t|; |X_t| > N) \leqslant \mathbb{E}[\mathbb{E}(|X||\mathscr{F}_t); |X_t| > N] = \mathbb{E}(|X|; |X_t| > N),$$

而

$$\mathbb{P}(|X_t| > N) \leqslant \frac{1}{N}\mathbb{E}|X_t| \leqslant \frac{1}{N}\mathbb{E}|X|.$$

实际上, 一致可积鞅也是 Doob 鞅(见习题). ▮

在这节的下面我们主要讨论离散时间鞅, 设 T 是离散时间集 $\{0, 1, 2, \cdots\}$, $(\Omega, \mathscr{F}, \mathbb{P})$ 是概率空间, $(\mathscr{F}_n : n \geqslant 0)$ 是流. 一个随机序列 $\{H_n : n \geqslant 0\}$ 称为是可预料的, 如果 H_0 是 $\mathscr{F}_0$ 可测的且对任何 $n \geqslant 1$, H_n 是 $\mathscr{F}_{n-1}$ 可测的. 设 X 是适应过程, H_n 是可预料过程, 定义

$$\begin{aligned}
&(H \cdot X)_0 := H_0 X_0, \\
&(H \cdot X)_n := (H \cdot X)_{n-1} + H_n(X_n - X_{n-1}), \ n \geqslant 1,
\end{aligned}$$

称为是过程 H 关于 X 的随机积分, 它是一般随机积分的离散形式. (为了在符号上区别乘积与随机积分, 除非必需, 我们写乘积时一般不用点.)

定理 3.2.1 设 X 是一个适应过程, H 是可预料过程使得 $H \cdot X$ 是可积的. 如果 X 是鞅, 那么过程 $H \cdot X$ 是鞅. 如果 X 是下鞅且 H 非负, 那么 $H \cdot X$ 是下鞅.

证明 显然 $H \cdot X$ 是适应的, 且对 $n \geqslant 1$,

$$\begin{aligned}\mathbb{E}[(H \cdot X)_n - (H \cdot X)_{n-1}|\mathscr{F}_{n-1}] &= \mathbb{E}(H_n(X_n - X_{n-1})|\mathscr{F}_{n-1}) \\ &= H_n\mathbb{E}(X_n - X_{n-1}|\mathscr{F}_{n-1}),\end{aligned}$$

因此 X 是鞅 (对应地,X 是下鞅及 H 的非负性) 蕴含着 $H \cdot X$ 是鞅 (对应地, 下鞅). □

称 H 是局部有界, 如果对任何 n, H_n 是有界的. 不难证明, 如果 X 可积, 且 H 局部有界, 那么 $H \cdot X$ 是可积的. 上面的定理以及习题中的 Doob 分解定理是随机分析中非常本质的结果, 有直观的解释. 设 X_n 是第 n 次赌博后某赌徒 A 的所有赌资, 则 $X_n - X_{n-1}$ 是 A 第 n 次赌博中输赢的数目, 另一个赌徒 B 赌 A 的运气, H_n 是乘子, 也就是 B 的策略. 但 B 也不可能预知下一局 A 的输赢, H_n 只能根据 $X_0, X_1, \cdots, X_{n-1}$ 的结果决定, 即 H_n 是 $\mathscr{F}_{n-1}$ 可测的, (从这个意义解释, 可预料也许应称为不可预料.) 定理指出 B 的运气不可能比 A 更好, 也不可能更坏.

作为一个特例, 设 τ 是停时, 定义 $H_n := 1_{\{\tau \geqslant n\}} = 1 - 1_{\{\tau \leqslant n-1\}}$, 则 $H_0 = 1$, 而 H_n 是 $\mathscr{F}_{n-1}$ 可测的且有界的. 随机积分

$$\begin{aligned}(H \cdot X)_n &= X_0 + \sum_{i=1}^{n} H_i(X_i - X_{i-1}) \\ &= X_n 1_{\{\tau \geqslant n\}} + X_{n-1} 1_{\{n > \tau \geqslant n-1\}} + \cdots + X_0 1_{\{\tau = 0\}} \\ &= X_{\tau \wedge n}.\end{aligned}$$

回忆上面定义的停止过程 $X_n^\tau := X_{\tau \wedge n}$, 我们得到下列推论.

推论 3.2.1 设 X 是一个鞅 (下鞅), τ 是一个停时, 则 τ 停止过程 X^τ 也是一个鞅 (下鞅).

下面的定理常称为 Doob 有界停止定理.

定理 3.2.2(Doob) 设 X 是一个鞅, σ, τ 是有界停时且 $\sigma \leqslant \tau$, 则 X_σ, X_τ 是可积的, $\mathbb{E}X_\sigma = \mathbb{E}X_\tau$ 且

$$\mathbb{E}(X_\tau|\mathscr{F}_\sigma) = X_\sigma \text{ a.s.}.$$

当 X 是下鞅时, 相应的结论成立.

证明 取 N, 使得 $\sigma, \tau < N$. 由推论3.2.1看出 X_σ, X_τ 是可积的, 且 $X_{\tau\wedge n} - X_{\sigma\wedge n}$, $n \in \mathsf{T}$ 是鞅, 因此对 $n \in \mathsf{T}$, $\mathbb{E}(X_{\tau\wedge n} - X_{\sigma\wedge n}) = 0$, 取 $n = N$, 得 $\mathbb{E}X_\sigma = \mathbb{E}X_\tau$.

显然 X_σ 是 $\mathscr{F}_\sigma$ 可测的, 对任何 $B \in \mathscr{F}_\sigma$, 令

$$\sigma_B := \sigma 1_B + N 1_{B^c}, \ \tau_B := \tau 1_B + N 1_{B^c}.$$

容易验证 σ_B 是停时, 另外由于 $\sigma \leqslant \tau$, 故 τ_B 也是停时. 因此

$$\mathbb{E}X_{\sigma_B} = \mathbb{E}X_{\tau_B},$$

或

$$\mathbb{E}(X_\sigma; B) + \mathbb{E}(X_N; B^c) = \mathbb{E}(X_\tau; B) + \mathbb{E}(X_N; B^c),$$

推出 $\mathbb{E}(X_\sigma; B) = \mathbb{E}(X_\tau; B)$, 即 $\mathbb{E}(X_\tau | \mathscr{F}_\sigma) = X_\sigma$ a.s.. 请读者自行验证当 X 是下鞅时的相应结论. □

上述定理也证明了下列推论.

推论 3.2.2 设 X 是一个可积适应过程, 则 X 是一个鞅当且仅当对任何有界停时 $\sigma \leqslant \tau$, 有 $\mathbb{E}X_\sigma = \mathbb{E}X_\tau$ 或等价地对任何有界停时 τ, $\mathbb{E}X_\tau = \mathbb{E}X_0$.

下面结果也是 Doob 停止定理的推论. 这个性质在后面将被使用.

推论 3.2.3 如果 X 是非负上鞅, 那么

$$\mathbb{P}(\{X_N > 0, \inf_{n \leqslant N} X_n = 0\}) = 0.$$

证明 固定 $N \geqslant 1$, $a > 0$, 则

$$\mathbb{P}(\{X_N \geqslant a, \inf_{n \leqslant N} X_n = 0\}) = 0.$$

事实上, 将其中的事件记为 A. 令

$$\tau := \min\{n \geqslant 0 : X_n = 0\};$$

$$\sigma := \min\{n \geqslant \tau : X_n \geqslant a\},$$

则 $A \subset \{\sigma \leqslant N\}$. 由 Doob 停止定理,

$$0 \geqslant \mathbb{E}(X_{\sigma\wedge N} - X_{\tau\wedge N})$$

$$
\begin{aligned}
&= \mathbb{E}(X_\sigma - X_\tau; \{\sigma \leqslant N\}) + \mathbb{E}(X_N - X_\tau; \{\sigma > N, \tau < N\}) \\
&\geqslant a\mathbb{P}(\sigma \leqslant N) + \mathbb{E}(X_N; \{\sigma > N, \tau < N\}) \\
&\geqslant a\mathbb{P}(\sigma \leqslant N).
\end{aligned}
$$

因此 $\mathbb{P}(\sigma \leqslant N) = 0$. 让 a 趋于 0 得所需结论. □

定理说明非负上鞅如果某处等于零就永远是零. 下面的例子说明 Doob 有界停止定理中有界性是重要的.

例 3.2.4 注意鞅只是说明每一局赌博是公平的, 但最后的结果可能仍然是不公平的. 一赌徒在开局时持有赌资 x, 如果赌徒可以无利率无限制借款且别人愿意在赌徒想玩的时候玩下去想停的时候停下来, 那么他总可以赢得任意多钱 $a > 0$ 后离开. 事实上, 设 X 是 $\mathbf{Z}$ 上 0 出发的简单随机游动, τ 是首次赢得至少 a 元钱的时刻. 由推论 3.2.1, X^τ 仍然是鞅, 也就是说赌博仍然是公平的. 但因 X 是常返的, 故 $\mathbb{P}(\tau < \infty) = 1$, 而 $X_\tau = a$, 因此 Doob 停止定理不成立. 也就是说有限时间内赌徒可以达到他的目的离开, 这样看来结果是不公平的.

另外一个有趣的策略是*加倍*. 赌徒为自己确定下列规则: 第 n 局下注为 a_n 直至第一次赢后离开. 这是一种典型的赌博策略, 如取 $a_n = a2^{n-1}$, 意味着赌徒的注每次加倍直至赢一次后离开. 记 τ 是赌徒第一次赢的局次, 即 $\tau = \inf\{n \geqslant 1 : X_n - X_{n-1} > 0\} = \inf\{n \geqslant 1 : \xi_n = 1\}$. 虽然 τ 不是首中时, 但显然

$$
\{\tau > n\} = \{\xi_1 = -1, \cdots, \xi_n = -1\} \in \mathscr{F}_n^0,
$$

故 τ 是停时. 令 $H_0 := 1$, $H_n := a_n \cdot 1_{\{n \leqslant \tau\}}$, 那么 (H_n) 是赌徒的下注策略. 显然 (H_n) 是可预料的. 在 n 次赌博后赌徒所拥有的赌资为随机积分

$$
(H \cdot X)_n = x + \sum_{i=1}^n H_i(X_i - X_{i-1}) = x + \sum_{i=1}^n H_i \xi_i.
$$

容易验证 $H \cdot X$ 是可积过程, 因此 $H \cdot X$ 是 $(\mathscr{F}_n^0)$ 鞅. 换句话说, 赌徒的运气实际上没有任何改变.

另外, 当 $\tau = n$ 时, $(H \cdot X)_n = x + a_n - \sum\limits_{i=1}^{n-1} a_i$, 那么赌徒在赢得一次后离开时的平均所得为

$$\begin{aligned}\mathbb{E}(H\cdot X)_\tau &= \sum_{n\geqslant 1}\mathbb{E}[(H\cdot X)_n;\{\tau=n\}]\\ &= x+\sum_{n\geqslant 1}\frac{1}{2^n}(2a_n-\sum_{k=1}^n a_k).\end{aligned}$$

如果 $\sum\limits_{n\geqslant 1}\dfrac{a_n}{2^n}$ 绝对收敛, 则交换和的次序得 $\mathbb{E}(H\cdot X)_\tau = x$. 但若加倍下注: $a_n = a2^{n-1}$, 则 $(H\cdot X)_\tau = a$, 即赌徒最终一定可以赢, Doob 停止定理也不成立.

为什么公平的游戏会产生不公平的结果呢? 原因就是不公平的规则: 赌徒可以无利率无限制借款且别人愿意在赌徒想玩的时候玩下去想停的时候停下来. 所以和一个有无限赌资的赌徒进行赌博依然是不公平的, 参考例 3.2.5. 在正规的赌场里一般都要限制最大的下注数或订其他规则以避免这种不公平的出现. ∎

鉴于 Doob 停止定理的重要性, 我们在此给出它的一个推广形式.

定理 3.2.3　设 $\{X_n : n\geqslant 0\}$ 是鞅, τ 是停时. 如果

(1) $\mathbb{P}(\tau<\infty)=1$;

(2) $\mathbb{E}|X_\tau|<\infty$;

(3) $\lim_n \mathbb{E}(X_n;\{\tau>n\})=0$;

那么 $\mathbb{E}X_\tau=\mathbb{E}X_0$.

证明　利用有界停止定理, 对任何 n,

$$\begin{aligned}\mathbb{E}X_0 &= \mathbb{E}X_{\tau\wedge n}\\ &= \mathbb{E}(X_\tau;\{\tau\leqslant n\})+\mathbb{E}(X_n;\{\tau>n\})\\ &\longrightarrow \mathbb{E}X_\tau.\end{aligned}$$

当 $n\to+\infty$ 时, 上面第 2 项因条件 (3) 趋于零, 第 1 项因条件 (1), (2) 及控制收敛定理也趋于 $\mathbb{E}X_\tau$. □

注　容易验证, 如果条件 (1) 满足且条件 (4) $\mathbb{E}(\sup_n|X_{\tau\wedge n}|)<\infty$ 满足, 则上面条件 (2),(3) 成立. 而显然 $\{X_{\tau\wedge n}: n\geqslant 0\}$ 的有界性蕴含着条件 (4) 成立.

下例是 Doob 停止定理的一个经典的应用.

例 3.2.5　设 $\{\xi_n\}$ 是一个 Bernoulli 随机序列. 取整数 $b>a>0$, 令 $X_0=a$, $X_n = X_0+\sum\limits_{k=1}^n\xi_k$, 它是从 a 出发的随机游动. 记 τ 为首次到达 $\{0,b\}$ 的时间, τ_0,τ_b

分别是首次到达 0 与 b 的时间. 容易验证 $\tau=\tau_0\wedge\tau_b$, $\mathbb{P}(\tau<+\infty)=1$. 直观地解释, 如用赢表示赌徒手中的赌资达到 b, 输表示赌徒输光, 那么 τ_0,τ_b 分别是赌徒首次输与赢的时间, τ 是首次输或赢的时间. 设赌徒无论输或赢都将立刻离开, 则 $\{\tau_0<\tau_b\}$ 表示赌徒最终输, 而 $\{\tau_0>\tau_b\}$ 表示赌徒最终赢. 记两者的概率分别为 p_0, p_b. 自然 $p_0+p_b=1$.

当 $p=q=\dfrac{1}{2}$ 时, X 是鞅. 对任何 $n\geqslant 0$, $0\leqslant X_{\tau\wedge n}\leqslant b$. 由定理 3.2.3 的注, $\mathbb{E}X_\tau=\mathbb{E}X_0=a$. 因 $\mathbb{E}X_\tau=0p_0+bp_b$, 故

$$p_0=\frac{b-a}{b},\ p_b=\frac{a}{b}.$$

因此赌资越多, 赢的概率越大.

当 $p>q$ 时, 因 $\mathbb{E}(q/p)^{\xi_n}=p+q=1$, 故不难验证 $\{(q/p)^{X_n}\}$ 是鞅, 那么 $\mathbb{E}(q/p)^{X_\tau}=(q/p)^a$. 同样

$$\mathbb{E}\left(\frac{q}{p}\right)^{X_\tau}=1\cdot p_0+\left(\frac{q}{p}\right)^b p_b.$$

因此

$$p_b=\frac{1-\left(\dfrac{q}{p}\right)^a}{1-\left(\dfrac{q}{p}\right)^b}.$$

(答案也可以用全概率公式列出差分方程解出.) ▮

现在设 X 是实值适应随机序列, 对 $-\infty<a<b<\infty$, 定义

$$\begin{aligned}
\tau_1 &:= \inf\{n\geqslant 0: X_n\leqslant a\};\\
\tau_2 &:= \inf\{n\geqslant \tau_1: X_n\geqslant b\};\\
&\cdots\ \cdots\\
\tau_{2k+1} &:= \inf\{n\geqslant \tau_{2k}: X_n\leqslant a\};\\
\tau_{2k+2} &:= \inf\{n\geqslant \tau_{2k+1}: X_n\geqslant b\};\\
&\cdots\ \cdots
\end{aligned}$$

则 $\{\tau_n : n \geqslant 1\}$ 是一个严格单调上升的停时序列, 对 $N \geqslant 1$, 令

$$U_N^X[a,b] := \max\{k : \tau_{2k} \leqslant N\},$$

随机变量 $U_N^X[a,b]$ 记录了随机序列 X 在时刻 0 与 N 之间从 a 下跳至 b 上的上窜次数. 下面是著名的 Doob 上窜不等式.

定理 3.2.4 (Doob)　设 X 是一个下鞅, 则对任何正整数 N, 常数 $a < b$,

$$\mathbb{E}U_N^X[a,b] \leqslant \frac{1}{b-a}[\mathbb{E}(X_N - a)^+ - \mathbb{E}(X_0 - a)^+].$$

证明　令 $Y_n := (X_n - a)^+$, 显然 $Y = (Y_n)$ 也是一个下鞅. 让 $\tau_1, \tau_2, \cdots$ 是将 $0, b-a, Y$ 分别取代 a, b, X 后如上定义的停时列, 自然 $U_N^X[a,b] = U_N^Y[0, b-a]$. 取 k 使 $2k-1 \geqslant N$, 那么 $\tau_{2k} \geqslant 2k-1 \geqslant N$, 因此

$$\begin{aligned}
Y_N - Y_0 &= \sum_{n=1}^{2k}(Y_{\tau_n \wedge N} - Y_{\tau_{n-1} \wedge N}) \\
&= \sum_{n=1}^{k}(Y_{\tau_{2n} \wedge N} - Y_{\tau_{2n-1} \wedge N}) + \sum_{n=0}^{k-1}(Y_{\tau_{2n+1} \wedge N} - Y_{\tau_{2n} \wedge N}) \\
&\geqslant (b-a)U_N^Y[0, b-a] + \sum_{n=0}^{k-1}(Y_{\tau_{2n+1} \wedge N} - Y_{\tau_{2n} \wedge N}).
\end{aligned}$$

而由定理 3.2.2,

$$\begin{aligned}
\mathbb{E}Y_N - \mathbb{E}Y_0 &\geqslant (b-a)\mathbb{E}U_N^X[a,b] + \sum_{n=0}^{k-1}(\mathbb{E}Y_{\tau_{2n+1} \wedge N} - \mathbb{E}Y_{\tau_{2n} \wedge N}) \\
&\geqslant (b-a)\mathbb{E}U_N^X[a,b].
\end{aligned}$$

□

Doob 上窜不等式是证明所有的鞅或下鞅收敛定理的基本工具.

定理 3.2.5　设 $\{X_n\}$ 是下鞅且 $K = \sup_n \mathbb{E}|X_n| < \infty$, 则 $X_n \to X$ a.s., 其中 X 是一个可积随机变量. 另外, 若 $\{X_n\}$ 是一个一致可积鞅, 则 $X_n \xrightarrow{L^1} X$ 且 $X_n = \mathbb{E}(X|\mathscr{F}_n)$.

证明　设 X^*, X_* 分别是 $\{X_n\}$ 的上极限与下极限. 显然

$$\{X^* > X_*\} = \bigcup_{a,b\in\mathbf{Q}} \{X_* < a < b < X^*\}.$$

由上窜不等式

$$\mathbb{E}U_N^X[a,b] \leqslant \frac{1}{b-a}(\mathbb{E}|X_N| + a) \leqslant \frac{K+a}{b-a}.$$

由单调收敛定理 $\mathbb{E}\lim_N U_N^X[a,b] < +\infty$, 因此 $\lim_N U_N^X[a,b] < +\infty$ a.s.. 但是 $\{X_* < a < b < X^*\} \subset \{\lim_N U_N^X[a,b] = +\infty\}$, 故有 $\mathbb{P}(\{X_* < a < b < X^*\}) = 0$, 推出 $X^* = X_*$ a.s., 极限的可积性由 Fatou 引理得到. 如果 $\{X_n\}$ 是一致可积鞅, 则由定理 1.4.7, $X_n \xrightarrow{L^1} X$ 且 $X_n = \lim_m \mathbb{E}(X_m|\mathscr{F}_n) = \mathbb{E}(X|\mathscr{F}_n)$. □

下面的鞅不等式也属于 Doob, 称为 Doob 鞅不等式.

定理 3.2.6 (Doob) 设 X 是一个下鞅.

(1) 对任何 $\lambda > 0$ 及正整数 N,

$$\lambda\mathbb{P}(\max_{0\leqslant n\leqslant N}|X_n| \geqslant \lambda) \leqslant 2\mathbb{E}X_N^+ - \mathbb{E}X_0;$$

(2) 设 X 是非负的, 则对任何 $p > 1$ 及正整数 N,

$$\mathbb{E}\max_{0\leqslant n\leqslant N} X_n^p \leqslant \left(\frac{p}{p-1}\right)^p \mathbb{E}X_N^p.$$

证明 (1) 令 $\tau := \min\{0 \leqslant n \leqslant N : X_n \geqslant \lambda\}$, 则 τ 是一个停时且 $\tau \leqslant N$, 故

$$\begin{aligned}\mathbb{E}X_N &\geqslant \mathbb{E}X_\tau \\ &= \mathbb{E}(X_\tau; \max_{0\leqslant n\leqslant N} X_n \geqslant \lambda) + \mathbb{E}(X_\tau; \max_{0\leqslant n\leqslant N} X_n < \lambda) \\ &\geqslant \lambda\mathbb{P}(\max_{0\leqslant n\leqslant N} X_n \geqslant \lambda) + \mathbb{E}(X_N; \max_{0\leqslant n\leqslant N} X_n < \lambda),\end{aligned}$$

因此

$$\lambda\mathbb{P}(\max_{0\leqslant n\leqslant N} X_n \geqslant \lambda) \leqslant \mathbb{E}(X_N; \max_{0\leqslant n\leqslant N} X_n \geqslant \lambda) \leqslant \mathbb{E}X_N^+.$$

另一方面, 如果令 $\tau := \min\{0 \leqslant n \leqslant N : X_n \leqslant -\lambda\}$, 则同样有

$$\mathbb{E}X_0 \leqslant \mathbb{E}X_\tau$$

$$= \mathbb{E}(X_\tau; \min_{0\leqslant n\leqslant N} X_n \leqslant -\lambda) + \mathbb{E}(X_\tau; \min_{0\leqslant n\leqslant N} X_n > -\lambda)$$

$$\leqslant -\lambda\mathbb{P}(\min_{0\leqslant n\leqslant N} X_n \leqslant -\lambda) + \mathbb{E}(X_N; \min_{0\leqslant n\leqslant N} X_n > -\lambda),$$

因此

$$\lambda\mathbb{P}(\min_{0\leqslant n\leqslant N} X_n \leqslant -\lambda) \leqslant \mathbb{E}(X_N; \min_{0\leqslant n\leqslant N} X_n > -\lambda) - \mathbb{E}X_0.$$

结合上述两个不等式, 将得到定理要证明的不等式.

(2) 令 $\xi := X_N$, $\eta := \max_{n\leqslant N} X_n$, $q := \dfrac{p}{p-1}$, 则由 (1) 的证明中知道 $t\mathbb{P}(\eta \geqslant t) \leqslant \mathbb{E}(\xi; \{\eta \geqslant t\})$, 再结合 Fubini 定理和 Hölder 不等式

$$\begin{aligned}
\mathbb{E}\eta^p &= \mathbb{E}\int_0^\eta pt^{p-1}\mathrm{d}t = \int_0^\infty pt^{p-1}\mathbb{P}(\eta \geqslant t)\mathrm{d}t \\
&\leqslant \int_0^\infty pt^{p-2}\mathbb{E}(\xi; \{\eta \geqslant t\})\mathrm{d}t \\
&\leqslant p\mathbb{E}\xi\int_0^\eta t^{p-2}\mathrm{d}t = \frac{p}{p-1}\mathbb{E}\xi\eta^{p-1} \\
&\leqslant q(\mathbb{E}\xi^p)^{\frac{1}{p}}(\mathbb{E}\eta^{(p-1)q})^{\frac{1}{q}} \\
&= q(\mathbb{E}\xi^p)^{\frac{1}{p}}(\mathbb{E}\eta^p)^{\frac{1}{q}}.
\end{aligned}$$

两边同除 $(\mathbb{E}\eta^p)^{\frac{1}{q}}$ 即得. □

定理 3.2.5 中下鞅的收敛形象地说是向右收敛, 在本节的最后, 我们将证明下鞅向左的收敛定理, 下面定理说明这样的收敛要容易得多.

定理 3.2.7 设 $X = (X_n)_{n\leqslant 0}$ 是关于流 $(\mathscr{F}_n)_{n\leqslant 0}$ 的下鞅且 $\inf_n \mathbb{E}X_n > -\infty$, 则

(1) X 是一致可积的;

(2) 当 $n \to -\infty$ 时, X_n 几乎处处且 L^1 收敛于一个可积随机变量 $X_{-\infty}$, 且对任何 n,

$$\mathbb{E}(X_n|\mathscr{F}_{-\infty}) \geqslant X_{-\infty},$$

其中 $\mathscr{F}_{-\infty} := \bigcap_{n=0}^{-\infty} \mathscr{F}_n$.

证明 设 $n \leqslant 0$, 因 $\mathbb{E}X_n \geqslant \mathbb{E}X_{n-1}$, 故 $\inf_n \mathbb{E}X_n > -\infty$ 蕴含着 $x = \lim_{n\to-\infty} \mathbb{E}X_n$ 存在且有限. 对给定 $\epsilon > 0$, 取 k 使得 $\mathbb{E}X_k - x < \epsilon$, 那么当 $n \leqslant k$ 时,

$$
\begin{aligned}
&\mathbb{E}(|X_n| : |X_n| > \lambda) = \mathbb{E}(X_n : X_n > \lambda) - \mathbb{E}(X_n : X_n < -\lambda) \\
= &\mathbb{E}(X_n : X_n > \lambda) + \mathbb{E}(X_n : X_n \geqslant -\lambda) - \mathbb{E}X_n \\
\leqslant &\mathbb{E}(X_k : X_n > \lambda) + \mathbb{E}(X_k : X_n \geqslant -\lambda) - \mathbb{E}X_k + \epsilon \\
\leqslant &\mathbb{E}(X_k : X_n > \lambda) + \mathbb{E}(-X_k : X_n < -\lambda) + \epsilon \\
\leqslant &\mathbb{E}(|X_k| : |X_n| > \lambda) + \epsilon.
\end{aligned}
$$

另外

$$
\begin{aligned}
\mathbb{P}(|X_n| > \lambda) &\leqslant \frac{1}{\lambda}\mathbb{E}|X_n| = \frac{1}{\lambda}\mathbb{E}(2X_n^+ - X_n) \\
&= \frac{1}{\lambda}(2\mathbb{E}X_n^+ - \mathbb{E}X_n) \leqslant \frac{1}{\lambda}(2\mathbb{E}X_0^+ - x),
\end{aligned}
$$

由此推出 X 是一致可积的.

(2) 类似于定理 3.2.5 的证明. 设 X^*, X_* 分别是当 $n \to -\infty$ 时 X_n 的上极限与下极限. 显然

$$
\{X^* > X_*\} = \bigcup_{a,b\in\mathbf{Q}} \{X_* < a < b < X^*\}.
$$

对正整数 N, 将上窜不等式用于下鞅 $\{X_{-N}, X_{-N+1}, \cdots, X_0\}$, 仍用 $U_N^X[a,b]$ 表示对 $[a,b]$ 的上窜次数, 则

$$
\mathbb{E}U_N^X[a,b] \leqslant \frac{1}{b-a}(\mathbb{E}|X_0| + a).
$$

因此同样有 $\lim_N U_N^X[a,b] < +\infty$ a.s.. 但是 $\{X_* < a < b < X^*\} \subset \{\lim_N U_N^X[a,b] = +\infty\}$, 故有 $\mathbb{P}(\{X_* < a < b < X^*\}) = 0$, 推出 $X^* = X_*$ a.s., 即 $\lim_{n\to-\infty} X_n(\omega)$ 几乎处处存在, 记极限为 $X_{-\infty}$. 再由 $\{X_n : n \geqslant 0\}$ 的一致可积性, X_n 也是 L^1 收敛于 $X_{-\infty}$, 因此 $X_{-\infty}$ 是可积的.

另外对 $A \in \mathscr{F}_{-\infty}$ 及 $m < n \leqslant 0$, 有 $\mathbb{E}(X_n; A) \geqslant \mathbb{E}(X_m; A)$, 让 $m \to -\infty$, $\mathbb{E}(X_n; A) \geqslant \lim_m \mathbb{E}(X_m, A) = \mathbb{E}(X_{-\infty}; A)$, 因此 $\mathbb{E}(X_n|\mathscr{F}_{-\infty}) \geqslant X_{-\infty}$. □

作为应用, 我们来证明 Kolmogorov 强大数定律. 它推广了 Borel 强大数律(参考 1.4.5).

定理 3.2.8(Kolmogorov)　设 $\{\xi_n\}$ 是独立同分布可积随机序列, 则

$$\frac{1}{n}(\xi_1+\xi_2+\cdots+\xi_n)$$

几乎处处且 L^1 收敛于 $\mathbb{E}\xi_1$.

证明　令 $X_{-n}:=\frac{1}{n}(\xi_1+\xi_2+\cdots+\xi_n)$, $\mathscr{F}_{-n}:=\sigma(X_{-k}:k\geqslant n)$, 那么 $(X_{-n},\mathscr{F}_{-n}:n\geqslant 1)$ 是鞅(见习题). 由上面的定理 3.2.7 推出 X_{-n} a.s. 且 L^1 收敛于一个可积随机变量 X.

不妨设 $\mathbb{E}\xi_1=0$, 则有 $\mathbb{E}X=0$. 让 ϕ 表示 ξ_1 的特征函数, 那么 ϕ 在 0 点可微且 $\phi'(0)=0$. 因此

$$\mathbb{E}\mathrm{e}^{itX}=\lim_n \mathbb{E}\mathrm{e}^{itX_{-n}}=\lim_n\left(\phi\left(\frac{t}{n}\right)\right)^n=1.$$

即 $X=0$ a.s.. □

习　题

1. 验证鞅定义下面所列的 3 个性质.
2. 设 $(Y_n:n\geqslant 1)$ 是一个有限状态 Markov 链, $\mathbf{P}$ 是转移矩阵, $\alpha:E\to\mathbf{R}$ 是 $\mathbf{P}$ 的从属于特征值 λ 的特征向量: $\mathbf{P}\alpha=\lambda\alpha$. 令 $X_n:=\lambda^{-n}\alpha(Y_n)$, 证明: $(X_n:n\geqslant 1)$ 是一个鞅.
3. 设 X 是平方可积鞅, $\mathbb{E}X_t\equiv 0$. 证明: X 是正交增量过程. 这说明鞅的性质介于独立增量与正交增量性质之间.
4. 设 X 是零均值平方可积的独立增量过程, 证明: 存在唯一的 T 上的递增函数 F, 使得 $(X_t^2-F(t))$ 是一个鞅.
5. 设 X 是一个右连续 $(\mathscr{F}_t)$ 下鞅, 如果存在一个适应连续递增过程 A, 使得 $A_0=0$ 且 $X-A$ 是 $(\mathscr{F}_t)$ 鞅, 那么称 X 是可补偿的, A 是 X 的补偿子.

   (1) 设 B 是 d-维 Brown 运动, 证明: (B_t^2) 是可补偿的;

   (2) 设 N 是参数为 λ 的 Poisson 过程, 问补偿 Poisson 过程的平方 $(N_t-\lambda t)^2$ 是否可补偿? N_t^2 是否可补偿? 如可以, 求补偿子.

6. 设 B 是 d-维 $(\mathscr{F}_t)$ Brown 运动, h 是 $\mathbf{R}^d$ 上的连续下调和函数. 证明: 若 $h\circ B_t$ 可积, 则它是下鞅. (提示: 用 Brown 运动的有限维分布计算 $\mathbb{E}(h\circ B_t|\mathscr{F}_s)$.)

7. (Wald 鞅) 设 $\{Y_n : n\geqslant 1\}$ 是独立同分布随机序列使得 $\phi(t) := \mathbb{E}e^{tY_n}$ 对某个 $t\neq 0$ 有限. 令

$$X_n := \phi(t)^{-n}\exp[t(Y_1+\cdots+Y_n)],$$

证明: $\{X_n : n\geqslant 1\}$ 是鞅.

8. 一个袋子中在时刻 0 有一个红球与一个白球. 随机地从袋子中取一个球, 然后将它放回并放入一个相同颜色的球, 无限地重复此过程. 记 X_n 为 n 次后袋中白球数与总球数之比, 证明: $\{X_n\}$ 是鞅.

9. 设 $\{Y_n : n\geqslant 1\}$ 是独立同分布随机序列, f_0, f_1 是两个概率密度函数, $f_0>0$. 令

$$X_n := \frac{f_1(Y_1)f_1(Y_2)\cdots f_1(Y_n)}{f_0(Y_1)f_0(Y_2)\cdots f_0(Y_n)}.$$

证明: 如果 f_0 是 Y_n 的密度函数, 那么 (X_n) 是鞅.

10. 设 $\{X_n : n\geqslant 0\}$ 是 $(\mathscr{F}_n)$ 适应的可积随机序列, 满足

$$\mathbb{E}(X_{n+1}|\mathscr{F}_n) = \alpha X_n + \beta X_{n-1},\ n\geqslant 1,$$

其中 $\alpha>0, \beta>0, \alpha+\beta=1$. 问 a 为何值时, 序列 $Y_0 := X_0$, $Y_n := aX_n + X_{n-1}$ 是 $(\mathscr{F}_n)$ 鞅?

11. 设 Markov 链 $\{X_n : n\geqslant 0\}$ 的状态空间为 $\{0,1,\cdots,N\}$, 转移概率为

$$p_{ij} = \binom{N}{j}\pi_i^j(1-\pi_i)^{N-j},\ 0\leqslant i,j\leqslant N,$$

其中 $\pi_i = \dfrac{1-e^{-2a\frac{i}{N}}}{1-e^{-2a}}$. 验证: $Z_n := e^{-2aX_n}$ 是鞅.

12. 设 $\{Y_n\}$ 是独立同分布正随机变量序列使得 $\mathbb{E}Y_n = 1$. 记 $X_n := Y_1Y_2\cdots Y_n$.

(1) 证明: (X_n) 是鞅且几乎处处收敛于一个随机变量 X;

(2) 设 Y_n 以概率 $\dfrac{1}{2}$ 分别取值 $\dfrac{1}{2}$ 与 $\dfrac{3}{2}$. 验证 $X=0$ a.s., 因此

$$\mathbb{E}\prod_{n\geqslant 1}Y_n \neq \prod_{n\geqslant 1}\mathbb{E}Y_n.$$

13. 设 $\{X_n : n \geqslant 1\}$ 是一个均值 0 且平方可积的鞅. 验证:

$$\mathbb{E}[(X_{n+k} - X_n)^2] = \sum_{i=1}^{k} \mathbb{E}[(X_{n+i} - X_{n+i-1})^2].$$

证明: 如果 $\sum_n \mathbb{E}[(X_{n+1} - X_n)^2] < \infty$, 那么 X_n 以概率 1 收敛.

14. (Doob 分解) 证明: 下鞅 (X_n) 可唯一分解为 $X_n = Y_n + Z_n$, 其中 (Y_n) 是鞅, (Z_n) 是从 0 出发的非负可预料的增过程: $0 = Z_1 \leqslant Z_2 \leqslant \cdots$.

15. (Riesz 分解) 证明: 上鞅 (X_n) 可分解为 $X_n = Y_n + Z_n$ 其中 Y 是鞅, Z 是位势(即 Z 是上鞅且 $\lim_n \mathbb{E}Z_n = 0$) 当且仅当 $\{\mathbb{E}X_n\}$ 有界. 这时此分解唯一.

16. 设 $\{X_n\}$ 是鞅且 $|X_0(\omega)|$, $|X_n(\omega) - X_{n-1}(\omega)|$ 被一个与 ω 及 n 无关的常数控制. 如果 τ 是一个具有限均值的停时, 证明: X_τ 可积且 $\mathbb{E}X_\tau = \mathbb{E}X_0$.

17. 设 $\{X_n : n \geqslant 1\}$ 是一个 Markov 链, 其转移矩阵是 $(p_{ij} : i, j \in E)$. E 上的非负函数 ϕ 称为是过分的, 如果 $\phi(i) \geqslant \sum_j p_{ij}\phi(j)$, $i \in E$. 用鞅理论证明如果 ϕ 是有界过分函数, 那么 $\phi(X_n)$ 以概率 1 收敛. 继而证明如果 X 是不可分并常返的, 那么 ϕ 是常数.

18. 设 $\{X_n\}$ 是 d-维对称随机游动. 证明: 当 $d = 1, 2$ 时, 有界过分函数是常数. 当 $d \geqslant 3$ 时, 存在非常数有界过分函数.

19. 考虑 2-维非负整数格点上的一个随机游动, 如果过程现在位于 (m, n), 那么下一步各以概率 $\dfrac{1}{2}$ 走到 $(m, n+1)$ 与 $(m+1, n)$. 设过程从 $(0, 0)$ 出发, $\varGamma$ 是一条从 x 轴的某格点联结相邻格点到 y 轴某格点的折线. 记 Y_1, Y_2 分别是过程碰到 $\varGamma$ 时向右与向上移动的步数. 证明: $\mathbb{E}Y_1 = \mathbb{E}Y_2$.

20. 设 $\{X_n : n \geqslant 0\}$ 是例 3.2.1 中定义的从 0 出发的随机游动, τ 是首次到达 $\{0, b\}$ 的时间. 证明: 当 $p = q$ 时, $X_n^2 - n$ 是鞅. 并由此求 $\mathbb{E}\tau$.

21. 设 $\{X_n : n \geqslant 0\}$ 是上面所说的随机游动, $p > q$. 取整数 $b > 0$, 令 $\sigma := \min\{n : X_n = b\}$. 求停时 σ 的母函数并由此计算 σ 的均值与方差.

22. 设 $\xi_1, \xi_2, \cdots, \xi_n, \cdots$ 是独立同分布可积随机变量. 对 $k \geqslant 1$, 令

$$X_{-k} := \frac{1}{k}(\xi_1 + \cdots + \xi_k),$$

$$\mathscr{F}_{-k} := \sigma(X_{-k}, X_{-k-1}, \cdots).$$

证明: $\{X_k : k \leqslant -1\}$ 关于 $(\mathscr{F}_k : k \leqslant -1)$ 是鞅.

§3.3 下鞅的正则化

在上一节鞅与下鞅的定义中没有涉及过程的轨道. 本节中我们将证明一个鞅或下鞅的轨道自动地具有很好的正则性, 即右连续并存在左极限.

设 $\mathsf{T} = [0,\infty)$, D 是 T 的一个可列稠子集, $(\Omega,\mathscr{F},\mathbb{P})$ 是完备概率空间, $(\mathscr{F}_t)_{t\in\mathsf{T}}$ 是概率空间上的流, 为方便计, 我们假设 $\mathscr{F}_0$ 中 含有 $\mathscr{F}$ 中的所有概率零的集合. 否则我们可加入这些集合. 这个假设不影响下面的结论, 因为所有的结果都是在几乎处处的意义下叙述的. 对 $K > 0$, 令 $D_K := [0,K] \cap (D \cup \{0,K\})$, 即 D_K 是包含 $0, K$ 及 其之间的 D 中点的全体.

现在设 $X = (X_t : t \in \mathsf{T})$ 是一个下鞅, 让 $F := \{0 = t_0 < t_1 < \cdots < t_N = K\}$ 是 D_K 的一个含有 $\{0,K\}$ 的有限子集, 对 $0 \leqslant n \leqslant N$, 令 $\mathscr{G}_n := \mathscr{F}_{t_n}$, $Y_n := X_{t_n}$, 则 $Y = (Y_n : 0 \leqslant n \leqslant N)$ 是一个下鞅, 由 Doob 的不等式,

$$\mathbb{E}U_N^Y[a,b] \leqslant \frac{1}{b-a}(\mathbb{E}(Y_N - a)^+ - \mathbb{E}(Y_0 - a)^+) \leqslant \frac{1}{b-a}\mathbb{E}(X_K - a)^+,$$

$$\lambda\mathbb{P}(\max_{0\leqslant n\leqslant N} |Y_n| > \lambda) \leqslant 2\mathbb{E}Y_N^+ - \mathbb{E}Y_0 = 2\mathbb{E}X_K^+ - \mathbb{E}X_0,$$

其中 $a < b$, $\lambda > 0$. 实际上 $U_N^Y[a,b]$ 是 X 限制在 F 上对 $[a,b]$ 的上窜次数, 我们记其为 $U_F^X[a,b]$, 再记

$$U_{D_K}^X[a,b] := \sup_{F\subset D_K} U_F^X[a,b],$$

它是 X 限制在 D_K 上对区间 $[a,b]$ 的上窜次数, 因为 D_K 可列, 故其有限子集全体是可列的, 因此 $U_{D_K}^X[a,b]$ 是一个非负可取无穷值的随机变量. 由上面的不等式推出

$$\mathbb{E}U_{D_K}^X[a,b] \leqslant \frac{1}{b-a}\mathbb{E}(X_K - a)^+,$$

$$\lambda\mathbb{P}(\max_{t\in D_K} |X_t| > \lambda) \leqslant 2\mathbb{E}X_K^+ - \mathbb{E}X_0.$$

下列定理是连续时间下鞅正则化的第一个基本定理.

定理 3.3.1　设 $X=(X_t:t\in\mathsf{T})$ 是一个下鞅, D 是 T 的一个可列稠子集, 则

(1) 对几乎所有的 $\omega\in\Omega$, 从 D 到 $\mathbf{R}$ 的映射 $t\mapsto X_t(\omega)$ 对任何 $K>0$ 在 D_K 上是有界的且在每个点 $t\in\mathsf{T}$ 有右极限

$$X_{t+}^D(\omega):=\lim_{s\in D,s\downarrow\downarrow t}X_s(\omega)$$

及左极限

$$X_{t-}^D(\omega):=\lim_{s\in D,s\uparrow\uparrow t}X_s(\omega).$$

(序列 $s_n\uparrow\uparrow t$ 表示对任何 $n\geqslant 1$, $s_n<t$ 且 $s_n\uparrow t$. $\downarrow\downarrow$ 类似理解.)

(2) 对所有 $t\in\mathsf{T}$, X_{t+}^D 是可积的且 $X_t\leqslant\mathbb{E}(X_{t+}^D|\mathscr{F}_t)$. 如果 $t\mapsto\mathbb{E}X_t$ 右连续, 则 $X_t=\mathbb{E}(X_{t+}^D|\mathscr{F}_t)$.

(3) 过程 $X_+^D=(X_{t+}^D:t\in\mathsf{T})$ 是关于右极限流 $(\mathscr{F}_{t+})$ 的右连续下鞅, 当 X 是鞅时, X_+^D 也是一个鞅.

证明　(1) 即 N_0 为 Ω 中使得映射 $t\mapsto X_t(\omega)$ 在某个 D_K 上无界或在某个点 t 处上述右极限或左极限不存在的 ω 全体. 我们无法直接验证 N_0 的可测性, 但是容易验证

$$N_0\subset\bigcup_{K\geqslant 1}\left[\{\max_{t\in D_K}|X_t|=\infty\}\cup N_K\right],$$

其中

$$N_K:=\bigcup_{a,b\in\mathbf{Q},a<b}\{U_{D_K}^X[a,b]=\infty\}.$$

而对任何固定的正整数 K 及 $a<b$, 由 Doob 的不等式,

$$\begin{aligned}\mathbb{P}(\{\max_{t\in D_K}|X_t|=\infty\})&=\lim_{\lambda\to\infty}\mathbb{P}(\{\max_{t\in D_K}|X_t|>\lambda\})\\&\leqslant\lim_{\lambda\to\infty}\frac{1}{\lambda}(2\mathbb{E}X_K^+-\mathbb{E}X_0)=0,\\\mathbb{P}(\{U_{D_K}^X[a,b]=\infty\})&=\lim_{N\to\infty}\mathbb{P}(\{U_{D_K}^X[a,b]\geqslant N\})\\&\leqslant\lim_{N\to\infty}\frac{1}{N}\mathbb{E}U_{D_K}^X[a,b]\\&\leqslant\lim_{N\to\infty}\frac{1}{N(b-a)}\mathbb{E}(X_K-a)^+=0.\end{aligned}$$

因概率空间是完备的, 故 $N_0 \in \mathscr{F}$ 且 $\mathbb{P}(N_0) = 0$.

(2) 对任何 $t \in \mathsf{T}$, 取 $\{s_n : n \geqslant 0\} \subset D$ 满足 $s_n \downarrow\downarrow t$. 对 $n \in \{0, -1, -2, \cdots\}$, 令 $\mathscr{G}_n := \mathscr{F}_{s_{-n}}$, $Y_n := X_{s_{-n}}$, 则 $Y = (Y_n : n = 0, -1, -2, \cdots)$ 是关于 $(\mathscr{G}_n : n = 0, -1, -2, \cdots)$ 的下鞅, 且

$$\mathbb{E}Y_n = \mathbb{E}X_{s_{-n}} \geqslant \mathbb{E}X_t.$$

由定理 3.2.6, Y 是一致可积的, 且 X_{s_n} 是 L^1-收敛于 X_{t+}^D, 因此 X_{t+}^D 是可积的, 又因为 $X_t \leqslant \mathbb{E}(X_{s_n}|\mathscr{F}_t)$, 由 L^1-收敛性推出 $X_t \leqslant \mathbb{E}(X_{t+}^D|\mathscr{F}_t)$. 如果 $t \mapsto \mathbb{E}X_t$ 右连续, 则 $\mathbb{E}X_t = \lim_n \mathbb{E}X_{s_n} = \mathbb{E}X_{t+}^D$, 那么非负随机变量 $\mathbb{E}(X_{t+}^D|\mathscr{F}_t) - X_t$ 的期望等于 $\mathbb{E}X_{t+}^D - \mathbb{E}X_t = 0$, 从而 $X_t = \mathbb{E}(X_{t+}^D|\mathscr{F}_t)$.

(3) 首先由于 $(\mathscr{F}_{t+})$ 是完备的, 容易看出 X_+^D 是 $(\mathscr{F}_{t+})$ 适应的可积过程, 另外对 $t > s$, 取 $s_n \in D$, $t > s_n \downarrow\downarrow s$, 还是由定理 3.2.6,

$$\mathbb{E}(X_t|\mathscr{F}_{s+}) \geqslant X_{s+}^D;$$

再取 $t_n \in D$, $t_n \downarrow\downarrow t$, 我们有 $\mathbb{E}(X_{t_n}|\mathscr{F}_{s+}) \geqslant X_{s+}^D$, 因此

$$\mathbb{E}(X_{t+}^D|\mathscr{F}_{s+}) = \lim_n \mathbb{E}(X_{t_n}|\mathscr{F}_{s+}) \geqslant X_{s+}^D,$$

即 X_+^D 是关于 $(\mathscr{F}_{t+})$ 的下鞅. 若 X 是鞅, 则 $\mathbb{E}X_t$ 是常数, 那么 $\mathbb{E}X_{t+}^D$ 也是常数, 因此 X_+^D 是鞅. □

注 . 上面右极限过程 X_+^D 的定义与可列稠集 D 有关, 但是两个不同可列稠集所定义的右极限过程是不可区分的(参见习题). 另外, 容易证明如果下鞅 X 是随机连续的, 那么 X_+^D 是 X 的一个修正.

上述定理有下面的重要推论. 定理中的两个结论是连续时间下鞅正则化的另两个基本定理.

定理 3.3.2 (1) 如果 X 是一个右连续的 $(\mathscr{F}_t)$ 下鞅, 则 X 也是一个 $(\mathscr{F}_{t+})$ 下鞅.

(2) 设流 $(\mathscr{F}_t)$ 满足通常条件. 如果下鞅 X 的均值 $t \mapsto \mathbb{E}X_t$ 是右连续的, 那么 X 有一个修正是右连左极的下鞅. 特别地, 鞅总有一个修正是右连左极的鞅.

证明 (1) 若 X 是右连续的, 则 X_+^D 与 D 的选取无关, 精确地说, 对不同 D 的选

取, X_+^D 是不可区分的且与 X 也是不可区分的, 因此 X 是 $(\mathscr{F}_{t+})$ 的下鞅.

(2) 当流与期望都是右连续时, 任取 T 的可列稠子集 D, 由定理 3.3.1,

$$X_t = \mathbb{E}(X_{t+}^D | \mathscr{F}_t) = X_{t+}^D, \ t \in \mathsf{T},$$

故 X_+^D 是 X 的修正, 而显然它是一个右连左极的下鞅. □

由这个定理, 以后我们说到右连续下鞅时, 总是可以假设对应的流是满足通常条件的. 另外如果流满足通常条件, 那么我们总可以假设鞅是右连左极的. 下面的定理是 Doob 有界停止定理的连续时间版本.

定理 3.3.3　设 X 是右连续 $(\mathscr{F}_t)$ 下鞅, τ, σ 是有界停时且 $\sigma \leqslant \tau$, 则 X_σ, X_τ 是可积的且

$$X_\sigma \leqslant \mathbb{E}(X_\tau | \mathscr{F}_\sigma), \text{ a.s..}$$

证明　对 $n \geqslant 0$, 令 $D_n := \left\{ \dfrac{k}{2^n} : k = 0, 1, 2, \cdots \right\}$ 及

$$\sigma_n(\omega) := \inf\{t \in D_n : \ t \geqslant \sigma(\omega)\}, \ \omega \in \Omega.$$

因 $D_n \subset D_{n+1}$, 故 σ_n 与 σ_{n+1} 是值域为 D_{n+1} 的关于流 $(\mathscr{F}_t : \ t \in D_{n+1})$ 的有界停时, 应用 Doob 离散时间有界停时定理于 $(\mathscr{F}_t : \ t \in D_{n+1})$ 下鞅 $(X_t : \ t \in D_{n+1})$, 得知 X_{σ_n} 与 $X_{\sigma_{n+1}}$ 是可积的且

$$X_{\sigma_{n+1}} \leqslant \mathbb{E}(X_{\sigma_n} | \mathscr{F}_{\sigma_{n+1}}).$$

对 $n = 0, -1, -2, \cdots$, 令 $Y_n := X_{\sigma_{-n}}$, $\mathscr{G}_n := \mathscr{F}_{\sigma_{-n}}$, 则 $(Y_n : n = 0, -1, -2, \cdots)$ 是关于 $(\mathscr{G}_n : n = 0, -1, -2, \cdots)$ 的下鞅, 且对任何 $n \leqslant 0$, $\mathbb{E}Y_n = \mathbb{E}X_{\sigma_{-n}} \geqslant \mathbb{E}X_0$, 由定理 3.2.6, $(X_{\sigma_n} : n = 0, 1, 2, \cdots)$ 是一致可积的, 且因为 X 是右连续的, 故当 $n \to \infty$ 时, X_{σ_n} 几乎处处收敛于 X_σ, 因而也 L^1-收敛于 X_σ, 因此 X_σ 是可积的, 同理 X_τ 也是可积的. 同样定义 $\tau_n := \inf\{t \in D_n : \ t \geqslant \tau\}$, 则 $\tau_n \geqslant \sigma_n$ 且都是有界停时, 对任何 $A \in \mathscr{F}_\sigma = \bigcap_n \mathscr{F}_{\sigma_n}$, 再应用 Doob 停止定理, 对任何 n,

$$\mathbb{E}(X_{\tau_n}; A) \geqslant \mathbb{E}(X_{\sigma_n}; A),$$

由 L^1-收敛性得 $\mathbb{E}(X_\tau; A) \geqslant \mathbb{E}(X_\sigma; A)$. □

推论 3.3.1　(1) 设 τ 是停时, X 是 $(\mathscr{F}_t)$ 右连续下鞅(或鞅), 则 X 的 τ 停止过程

$X^\tau = (X_{t\wedge\tau}:\ t \in \mathsf{T})$ 也是关于 $(\mathscr{F}_t)$ 的下鞅(鞅).

(2) $(\mathscr{F}_t)$ 适应的右连续实值可积过程 X 是鞅当且仅当对任何有界停时 τ 有 $\mathbb{E}X_\tau = \mathbb{E}X_0$.

证明 (1) 因 $t\wedge\tau$ 是有界停时, 由上述定理, X^τ 是 $(\mathscr{F}_t)$ 适应的可积过程, 并且对 $t > s$, 再应用上述定理及由定理 3.1.2,

$$\begin{aligned}\mathbb{E}(X_{t\wedge\tau}|\mathscr{F}_s) &= \mathbb{E}(\mathbb{E}(X_{t\wedge\tau}|\mathscr{F}_{t\wedge\tau})|\mathscr{F}_s)\\ &= \mathbb{E}(X_{t\wedge\tau}|\mathscr{F}_{t\wedge\tau\wedge s}) = \mathbb{E}(X_{t\wedge\tau}|\mathscr{F}_{s\wedge\tau})\\ &\geqslant X_{s\wedge\tau}.\end{aligned}$$

故 X^τ 也是下鞅. (2) 可由推论 3.2.2 直接推出. □

下面定理类似于推论 3.2.3. 说明一个非负右连续上鞅一旦在某时刻等于零就永远是零.

定理 3.3.4 设 X 是非负右连续上鞅, $\tau := \inf\{t:\ X_t \cdot X_{t-} = 0\}$, 则对 a.s. $\omega \in \Omega$, $t > \tau(\omega)$ 时, $X_t(\omega) = 0$.

证明 令 $\tau_n := \inf\left\{t \geqslant 0 : X_t < \dfrac{1}{n}\right\}$, 那么 τ_n 是停时且当 $\tau_n < +\infty$ 时, $X_{\tau_n} \leqslant \dfrac{1}{n}$. 由定理 3.3.3, $\mathbb{E}(X_t|\mathscr{F}_{\tau_n}) \leqslant X_{t\wedge\tau_n}$, 因此

$$\mathbb{E}(X_t; \tau_n \leqslant t) \leqslant \mathbb{E}(X_{t\wedge\tau_n}; \tau_n \leqslant t) = \mathbb{E}(X_{\tau_n}; \tau_n \leqslant t) \leqslant \frac{1}{n}.$$

而 $\tau_n \uparrow \tau$, 故 $\mathbb{E}(X_t; \tau \leqslant t) = 0$, 再由 X 的右连续性推出定理结论. □

另外, 定理 3.2.3 中的两个下鞅不等式对限制在可列稠子集上成立, 由下鞅的右连续性 Doob 的不等式可推广到连续时间下鞅.

定理 3.3.5 设 X 是一个右连续下鞅, 则

(1) 对任何 $\lambda > 0$ 及 $N > 0$,

$$\lambda\mathbb{P}(\max_{0\leqslant t\leqslant N} |X_t| \geqslant \lambda) \leqslant 2\mathbb{E}X_N^+ - \mathbb{E}X_0;$$

(2) 若 X 是非负的, 则对任何 $p > 1$ 及 $N > 0$,

$$\mathbb{E}\max_{0\leqslant t\leqslant N} X_t^p \leqslant \left(\frac{p}{p-1}\right)^p \mathbb{E}X_N^p.$$

特别地, 如果 X 是鞅, 那么 $\mathbb{E}\max_{0\leqslant t\leqslant N} X_t^2 \leqslant 4\mathbb{E}X_N^2$.

证明留给读者作为习题. 最后, 我们将要介绍的局部鞅的概念, 由推论 3.3.1, 局部鞅的引入显得非常自然, 它在随后几节中将会经常用到.

定义 3.3.1　固定一个满足通常条件的流 $(\mathscr{F}_t)$. 右连续适应过程 M 称为是一个局部鞅, 如果存在一个单增趋于无穷的停时列 $\{T_n\}$, 使得对任何 n, 过程 $M^{T_n}1_{\{T_n>0\}}$ 是一个鞅. 这时, $\{T_n\}$ 称为是 M 的局部化序列.

因为, $M^{T_n} = M^{T_n}1_{\{T_n>0\}} + M_0 1_{\{T_n=0\}}$, 故如果 M_0 本身是可积的, 那么定义等价于一个看上去更自然的说法: 对任何 n, M^{T_n} 是一个鞅. 用以上定义的一个重要理由是包含一些平凡的情形, 例如恒等于一个 $\mathscr{F}_0$ 可测随机变量的实值过程在以上定义下是局部鞅(见习题). 实际上, 对任何连续局部鞅 M, 总是存在一个局部化序列 $\{T_n\}$, 使得 $M^{T_n}1_{\{T_n>0\}}$ 是一个有界连续鞅. 以后连续局部鞅的局部化序列总是指这样一个序列. 显然鞅一定是局部鞅. 反过来, 局部鞅不一定是鞅. 但是由 Fatou 引理, 可积非负局部鞅是上鞅. 由有界收敛定理, 有界局部鞅是鞅. 局部鞅是鞅的一个非常有用的推广, 在后面几节中将会看到这一点. 在这里我们给一个可积的非负局部鞅但不是鞅的例子.

例 3.3.1　设 $B = (B_t)$ 是 $\mathbf{R}^3$ 上的 Brown 运动, $h(x) := \dfrac{1}{|x|}$, $x \in \mathbf{R}^3 \setminus \{0\}$, 则 $h(B) = \{h(B_t)\}$ 是局部鞅, 但它不是鞅. 对任何 $k \geqslant 1$, 取

$$D_k := \left\{x \in \mathbf{R}^3 : |x| \leqslant \frac{1}{k}\right\},$$

则 h 在 $\mathbf{R}^3 \setminus \{0\}$ 上调和, 那么 $h(B)$ 在概率 $\mathbb{P}^{x_0}$, $x_0 \neq 0$ 是局部鞅而不是鞅.

首先由定理 4.4.3, 单点是极集, 因此过程 $h(B)$ 的状态空间是 $\mathbf{R}^3$. 令 T_k 是 D_k^c 的首中时, 那么对任何 $x_0 \neq 0$, 停止过程 $\{h(B_{t\wedge T_k}\}_{t\geqslant 0}$ 关于概率 $\mathbb{P}^{x_0}$ 是有界鞅(见§5节中的证明). 而仍然由定理 4.4.3 $T_k \uparrow T_{\{0\}} = \infty$, 故推出 $h(B)$ 是局部鞅. 另外我们知道, 如果 $h(B)$ 是鞅, 则对任何 $t > 0$,

$$\mathbb{E}^{x_0}h(B_t) = \mathbb{E}^{x_0}h(B_0) = h(x_0) = \frac{1}{|x_0|}.$$

而对任何 $N > |x_0|$,

$$
\begin{aligned}
\mathbb{E}^{x_0}h(B_t) &= \int_{\mathbf{R}^3} \frac{1}{(2\pi t)^{\frac{3}{2}}} \frac{1}{|y|} \mathrm{e}^{-\frac{|y-x_0|^2}{2t}} \mathrm{d}y \\
&\leqslant \int_{|y|\leqslant N} \frac{1}{(2\pi t)^{\frac{3}{2}}} \frac{1}{|y|} \mathrm{e}^{-\frac{|y-x_0|^2}{2t}} \mathrm{d}y + \int_{|y|>N} \frac{1}{N} \frac{1}{(2\pi t)^{\frac{3}{2}}} \mathrm{e}^{-\frac{|y-x_0|^2}{2t}} \mathrm{d}y \\
&\leqslant \frac{C_1}{(2\pi t)^{\frac{3}{2}}} + \frac{C_2}{N},
\end{aligned}
$$

其中常数 C_1 与 t 无关, C_2 与 N, t 都无关, 因此 $\lim_{t\to\infty} \mathbb{E}^{x_0}h(B_t) = 0$, 故 $h(B)$不是鞅. ∎

习　题

1. 设 D_1, D_2 是 T 的两个可列稠子集, X^1, X^2 是分别由 D_1, D_2 按定理 3.3.1 中方法定义的右极限过程. 证明: X^1 与 X^2 是不可区分的.
2. (下鞅收敛定理) 设 (X_t) 是 $(\mathscr{F}_t)$ 的右连续下鞅且 $\sup_{t\in\mathsf{T}} \mathbb{E}X_t^+ < \infty$. 证明: $\lim_{t\to\infty} X_t$ 几乎处处收敛且极限可积.
3. 设 Z 是可积随机变量. 证明: 对任何 $t_n \downarrow 0$,
$$\lim_n \mathbb{E}(Z|\mathscr{F}_{t_n}) = \mathbb{E}(Z|\mathscr{F}_{0+}).$$
4. 设 (X_t) 是一个关于 $(\mathscr{F}_t)$ 的右连续非负上鞅. 证明: 当 $t\to\infty$ 时, X_t 几乎处处收敛于一个可积随机变量(记为 X_∞) 且 $(X_t : 0\leqslant t\leqslant +\infty)$ 是 $(\mathscr{F}_t)$ 上鞅.
5. (Riesz 分解) 一个右连续的 $(\mathscr{F}_t)$ 非负上鞅 (Z_t) 称为位势, 如果 $\lim_{t\to\infty} \mathbb{E}Z_t = 0$. 证明: 右连续一致可积上鞅 (X_t) 有 Riesz 分解 $X_t = M_t + Z_t$, 其中 (M_t) 是一个右连续一致可积的 $(\mathscr{F}_t)$ 鞅, (Z_t) 是一个位势.
6. 证明: 一个鞅是一致可积鞅当且仅当它是右闭鞅 (Doob 鞅).
7. 设 $(X^{(n)})$ 是一个右连续上鞅列且对任何 t, $X_t^{(n)}$ 关于 n 递增. 证明:极限 $X_t := \sup_n X_t^{(n)}$ 右连续且左极限存在.
8. 设 (N_t) 是参数为 λ 的 Poisson 过程, 对 $u\in\mathbf{C}$, 定义
$$X_t := \exp[iuN_t - \lambda t(\mathrm{e}^{iu} - 1)];\ \ t\geqslant 0.$$

(1) 证明: $(\mathrm{Re}(X_t))$, $(\mathrm{Im}(X_t))$ 是鞅;

(2) 当 $u=-i$ 时, 上面的鞅是否一致可积.

9. 证明: M 是局部鞅当且仅当存在停时列 $T_n \uparrow +\infty$, 使得 M_0 在 $\{T_n>0\}$ 上可积且 $M^{T_n}-M_0$ 是鞅.

10. 设 M_0 可积. 证明: M 是局部鞅当且仅当存在停时列 $T_n \uparrow +\infty$, 使得对任何 n, M^{T_n} 是鞅.

11. 证明:

(1) 设 ξ 是 $\mathscr{F}_0$ 可测随机变量, X 是局部鞅, 那么 $\xi\cdot X$ 也是局部鞅;

(2) 非负局部鞅是上鞅;

(3) 局部鞅全体是个线性空间;

(4) 局部鞅的停止过程仍是局部鞅.

12. 设 M 是连续局部鞅. 证明: 存在递增停时列 $\{T_n\}$, 使得对任何 n, $M^{T_n}1_{\{T_n>0\}}$ 是有界连续鞅.

13. 右连续适应过程 M 是局部鞅当且仅当存在停时列 $\tau_n \uparrow +\infty$, 使得过程 $M^{\tau_n}\cdot 1_{\{\tau_n>0\}}$ 是一致可积鞅.

14. 证明 M 是局部鞅当且仅当存在停时列 $T_n \uparrow +\infty$, 且对任何 n, M^{T_n} 是局部鞅.

15. 设 B 是标准 Brown 运动, $a,b>0$, 令 $T:=\inf\{t>0:\ B_t>a+bt\}$, 求 $\mathbb{P}(T<+\infty)$. (提示: 参考文献 [12].)

§3.4 随机积分与 Itô 公式

在本节中, 设 $\mathsf{T}=[0,\infty)$, $(\mathscr{F}_t)$ 是概率空间上的一个满足通常条件的流. 设 $\mathscr{M}^2$ 是平方可积鞅全体, 而 $\mathscr{M}_c^2$ 是连续平方可积鞅全体, 当然 $\mathscr{M}_c^2$ 是 $\mathscr{M}^2$ 的线性子空间. 对 $M\in\mathscr{M}^2$, 定义

$$\|M\|:=\sum_{n\geqslant 0}2^{-n}(1\wedge\sqrt{\mathbb{E}M_n^2}).$$

显然它诱导 $\mathscr{M}^2$ 上的度量. 我们先证明一个完备性定理.

定理 3.4.1 $(\mathscr{M}^2,\|\cdot\|)$ 是完备的, $\mathscr{M}_c^2$ 是闭子空间.

证明　设 $M^{(n)}$ 是 $\mathscr{M}^2$ 中的一个 Cauchy 列, 则显然对任何 $t \geqslant 0$, 存在 M_t, 使得 $M_t^{(n)}$ 平方 (L^2-)收敛于 M_t. 容易验证 $M = (M_t) \in \mathscr{M}^2$ 且 $M^{(n)}$ 以 $\|\cdot\|$ 收敛于 M, 因此 $\mathscr{M}^2$ 是完备的. 进一步, 由 Doob 不等式(定理 3.3.5(2)), 对任何 $t > 0$, 有

$$\mathbb{E} \max_{0 \leqslant s \leqslant t} (M_s^{(n)} - M_s)^2 \leqslant 4\mathbb{E}(M_t^{(n)} - M_t)^2.$$

由此推出除去一个零概率集外, $t \mapsto M_t^{(n)}$ 在 $[0, \infty)$ 上紧一致收敛(至少沿一个子列)于 $t \mapsto M_t$. 因此 $M^{(n)} \in \mathscr{M}_c^2$ 蕴含着 $M \in \mathscr{M}_c^2$, 即 $\mathscr{M}_c^2$ 是闭的. □

一个适应右连续随机过程 A 称为是增过程 (或有界变差过程), 如果 $A_0 = 0$ a.s. 且对几乎所有的 $\omega \in \Omega$, 轨道 $t \mapsto A_t(\omega)$ 是单调上升的 (对应地, 在有限区间上是有界变差的). 首先我们证明像 Brown 运动一样, 一个非常值连续鞅不可能是有界变差过程.

定理 3.4.2　一个连续局部鞅是有界变差过程当且仅当它恒等于零.

证明　设 M 是一个具有有界变差的连续局部鞅. 记 V 是 M 的全变差过程, 定义 $T_n := \inf\{t: \ V_t \geqslant n\}$, 则 $\{T_n\}$ 是一个趋于无穷的单增停时列, 而且停止过程 M^{T_n} 是一个具有有界的全变差过程的连续局部鞅. 因此下面我们不妨设 M 的全变差过程及 M 本身被常数 K 控制. 对任何 $t \geqslant 0$ 及 $[0, t]$ 上的任何分划 $\varDelta = \{t_i\}$, 因为 $M_0 = 0$, 故有

$$\mathbb{E}M_t^2 = \mathbb{E} \sum_i (M_{t_{i+1}}^2 - M_{t_i}^2) = \mathbb{E} \sum_i (M_{t_{i+1}} - M_{t_i})^2 \leqslant K \cdot \mathbb{E} \sup_i |M_{t_{i+1}} - M_{t_i}|.$$

由 M 的连续性, 当 $|\varDelta| \to 0$ 时, $\sup_i |M_{t_{i+1}} - M_{t_i}| \to 0$ a.s., 再由控制收敛定理, $\mathbb{E} \sup_i |M_{t_{i+1}} - M_{t_i}| \to 0$, 故 $\mathbb{E}M_t^2 = 0$, 从而 M 恒等于零. □

虽然连续局部鞅一般没有有界一次变差, 但它却有有界二次变差, 而且这个性质使得我们可以定义关于连续鞅的随机积分. 对于任意的适应右连续随机过程 X, 如果存在一个右连续随机过程 A, 使得当 $[0, \infty)$ 上的任何变细的分划 $\varDelta := \{t_i\}$ 的长度趋于零时, 对任何 $t > 0$, 平方和过程

$$T_t^{\Delta}(X) := \sum_i (X_{t_{i+1} \wedge t} - X_{t_i \wedge t})^2$$

依概率收敛于 A_t, 则称 X 具有有界的二次变差. 过程 A 称为是 X 的二次变差过

程, 写为 $\langle X\rangle$. (当我们说 $\Delta=\{t_i\}$ 是 $[0,\infty)$ 上的分划时, 总约定在任何 t 前只有有限多分划点, 另外 $t_i\to\infty$.) 显然一个连续有界变差过程具有有界二次变差, 其二次变差过程是零过程. 下面我们证明连续局部鞅具有有界的二次变差.

定理 3.4.3 设 M 是一个连续局部鞅, 则

(1) M 具有有界二次变差;

(2) 其二次变差过程 $\langle M\rangle$ 就是使得 M^2-A 成为连续局部鞅的唯一的连续增过程 A.

证明 唯一性由定理 3.4.2 立即推出. 现在证明二次变差过程的存在性. 在本定理证明中, 过程 X 局部有界是指对任何 $t\geq 0$, X_t 有界. 容易验证如果 M 有界, 则 $T^\Delta(M)$ 局部有界; 如果 M 局部有界, 因为 M 是鞅, 故对任何 $t>s$, 控制 M_t 的常数一定控制 M_s, 因此 $T_t^\Delta(M)$ 仍然局部有界. 先设 M 是有界连续鞅, 对任何 $t>s\geqslant 0$, 存在 k, 使 $t_k<s\leqslant t_{k+1}$, 故

$$
\begin{aligned}
&\mathbb{E}(T_t^\Delta(M)-T_s^\Delta(M)|\mathscr{F}_s)\\
&=\mathbb{E}\left[\sum_{i>k}(M_{t_{i+1}\wedge t}-M_{t_i\wedge t})^2|\mathscr{F}_s\right]+\mathbb{E}\left[(M_{t_{k+1}\wedge t}-M_{t_k})^2-(M_s-M_{t_k})^2|\mathscr{F}_s\right]\\
&=\mathbb{E}\left(\sum_{i>k}(M_{t_{i+1}\wedge t}-M_{t_i\wedge t})^2|\mathscr{F}_s\right)+\mathbb{E}((M_{t_{k+1}\wedge t}-M_s)^2|\mathscr{F}_s)\\
&=\sum_i\mathbb{E}((M_{(t_{i+1}\wedge t)\vee s}-M_{(t_i\wedge t)\vee s})^2|\mathscr{F}_s)\\
&=\sum_i\mathbb{E}(M^2_{(t_{i+1}\wedge t)\vee s}-M^2_{(t_i\wedge t)\vee s}|\mathscr{F}_s)\\
&=\mathbb{E}(M_t^2-M_s^2|\mathscr{F}_s),
\end{aligned}
$$

即 $M^2-T^\Delta(M)$ 是一个局部有界连续鞅. 现取 $[0,\infty)$ 的任意两个分划 Δ 与 Δ', 用 Δ'' 表示两者合并后的分划, 则因 $T^\Delta(M)-T^{\Delta'}(M)$ 是局部有界连续鞅, 故由以上论证, 过程

$$
(T_t^\Delta(M)-T_t^{\Delta'}(M))^2-T_t^{\Delta''}(T^\Delta(M)-T^{\Delta'}(M)),\ t\geqslant 0
$$

是局部有界连续鞅且容易看出

$$T_t^{\Delta''}(T^{\Delta}(M)-T^{\Delta'}(M)) \leqslant 2T_t^{\Delta''}(T^{\Delta}(M))+2T_t^{\Delta''}(T^{\Delta'}(M)).$$

下面我们验证当分划的长度趋于零时, 等式右边 L^1-收敛于零. 事实上, 设 $\{t_i'\}$, $\{s_i'\}$ 分别是 Δ, Δ'' 的分划点, 再记 $t_i := t_i' \wedge t$, $s_i := s_i' \wedge t$, 这时对任何 k, 存在唯一的 l 使得 $t_l \leqslant s_k \leqslant s_{k+1} \leqslant t_{l+1}$, 因此

$$\begin{aligned}
T_{s_{k+1}}^{\Delta}(M)-T_{s_k}^{\Delta}(M) &= (M_{s_{k+1}}-M_{t_l})^2-(M_{s_k}-M_{t_l})^2 \\
&= (M_{s_{k+1}}-M_{s_k})(M_{s_{k+1}}+M_{s_k}-2M_{t_l}),
\end{aligned}$$

从而

$$\begin{aligned}
T_t^{\Delta''}(T^{\Delta}(M)) &= \sum_k (T_{s_{k+1}}^{\Delta}(M)-T_{s_k}^{\Delta}(M))^2 \\
&\leqslant T_t^{\Delta''}(M)\cdot \sup_k (M_{s_{k+1}}+M_{s_k}-2M_{t_l})^2.
\end{aligned}$$

由 Cauchy-Schwarz 不等式,

$$[\mathbb{E}T_t^{\Delta''}(T^{\Delta}(M))]^2 \leqslant \mathbb{E}[T_t^{\Delta''}(M)]^2\cdot \mathbb{E}\sup_k (M_{s_{k+1}}+M_{s_k}-2M_{t_l})^4.$$

现在来估计 $\mathbb{E}(T_t^{\Delta}(M))^2$. 用 C 表示 M 的一个界, 记 $s_i := t_i \wedge t$,

$$\begin{aligned}
&\mathbb{E}(T_t^{\Delta}(M))^2 = \mathbb{E}\left[\sum_i (M_{s_{i+1}}-M_{s_i})^2\right]^2 \\
&= \mathbb{E}\sum_i (M_{s_{i+1}}-M_{s_i})^4 + 2\mathbb{E}\sum_{i>j}(M_{s_{i+1}}-M_{s_i})^2(M_{s_{j+1}}-M_{s_j})^2 \\
&= \mathbb{E}\sum_i (M_{s_{i+1}}-M_{s_i})^4 + 2\mathbb{E}\sum_{i>j}(M_{s_{i+1}}^2-M_{s_i}^2)(M_{s_{j+1}}-M_{s_j})^2 \\
&= \mathbb{E}\sum_i (M_{s_{i+1}}-M_{s_i})^4 + 2\mathbb{E}\sum_{j}(M_t^2-M_{s_{j+1}}^2)(M_{s_{j+1}}-M_{s_j})^2 \\
&= \mathbb{E}\sum_i (M_{s_{i+1}}-M_{s_i})^2[(M_{s_{i+1}}-M_{s_i})^2+2(M_t^2-M_{s_{i+1}}^2)] \\
&\leqslant 8C^2\mathbb{E}T_t^{\Delta}(M) = 8C^2\mathbb{E}(M_t^2-M_0^2) \leqslant 16C^4,
\end{aligned}$$

即它被一个与分划无关的常数所控制.

因此由有界收敛定理推出 $\lim_{|\Delta|\to 0}\mathbb{E}T_t^{\Delta''}(T^{\Delta}(M))=0$, 故

$$\lim_{|\Delta|,|\Delta'|\to 0}\mathbb{E}(T_t^{\Delta}(M)-T_t^{\Delta'}(M))^2=0,$$

取长度趋于零的一个分划列 $\{\Delta_n\}$, 则 $\{M^2-T^{\Delta_n}(M)\}$ 是 $\mathscr{M}_c^2$ 中的一个 Cauchy 列, 由定理 3.4.1, 它在 $\mathscr{M}_c^2$ 中有极限. 因此存在一个连续过程 A, 使得对任何 $t\geqslant 0$, $T_t^{\Delta_n}(M)$ 是 L^2 故而也是依概率收敛于 A_t 且 M^2-A 是一个鞅. 因 $T_0^{\Delta_n}(M)=0$ a.s., 故 $A_0=0$ a.s.. 另外对任何 $t>s$, $s,t\in\bigcup_n\Delta_n$, 存在充分大 n, 使得

$$T_t^{\Delta_n}(M)\geqslant T_s^{\Delta_n}(M),\ \text{a.s.},$$

因此 $A_t\geqslant A_s$ a.s.. 由 A 的连续性推出 A 是一个增过程, 由唯一性, A 与 Δ_n 的选择无关, 故定理对有界连续鞅成立.

由唯一性容易验证, 如果 M 有界, T 是停时, 则 $\langle M^T\rangle=\langle M\rangle^T$. 现在设 M 是一个局部鞅, 取其一个局部化序列 $\{T_n\}$, 则由上面的说明得,

$$\langle M^{T_n}\rangle=\langle (M^{T_{n+1}})^{T_n}\rangle=\langle M^{T_{n+1}}\rangle^{T_n}.$$

对 $t\leqslant T_n$, 定义 $\langle M\rangle_t:=\langle M^{T_n}\rangle_t$, (实际上是暂时借用记号 $\langle M\rangle$.) 上式说明定义无歧义, 容易验证 $\langle M\rangle$ 是一个从零出发的连续增过程且 $M^2-\langle M\rangle$ 是一个连续局部鞅(作习题).

最后我们证明 $\langle M\rangle$ 是 M 的二次变差过程(说明上面借用的记号是合理的). 因为 M 是连续局部鞅, 故存在局部化停时列 $\{T_n\}$ 使得 M^{T_n} 是有界连续鞅, 那么 $\langle M\rangle^{T_n}=\langle M^{T_n}\rangle$. 对任何 $t\geqslant 0$, 当 $k\to\infty$ 时,

$$T_t^{\Delta_k}(M^{T_n})\xrightarrow{L^2}\langle M\rangle_t^{T_n}.$$

因此在 $\{t<T_n\}$ 上 $T_t^{\Delta_k}(M)\xrightarrow{L^2}\langle M\rangle_t$. 现在任取 $\epsilon>0$, 对任何 $n\geqslant 1$, 有

$$\begin{aligned}&\limsup_{k\to\infty}\mathbb{P}(|T_t^{\Delta_k}(M)-\langle M\rangle_t|>\epsilon)\\ \leqslant\ &\lim_k\mathbb{P}(|T_t^{\Delta_k}(M)-\langle M\rangle_t|>\epsilon,t<T_n)+\mathbb{P}(t\geqslant T_n)=\mathbb{P}(t\geqslant T_n),\end{aligned}$$

而 $\lim_n\mathbb{P}(t\geqslant T_n)=0$, 即 $T_t^{\Delta_k}(M)\xrightarrow{p}\langle M\rangle_t$. 完成了证明. □

由定理看出如果连续局部鞅 M 的二次变差过程 $\langle M\rangle$ 恒等于 0, 那么 $M\equiv M_0$. 对任何连续局部鞅 M,N 定义

$$\langle M,N\rangle := \frac{1}{4}(\langle M+N\rangle - \langle M-N\rangle),$$

称它是 M,N 的协变差过程. 自然地, 对任何 $t\geqslant 0$, 当分划 $\{t_i\}$ 趋于零时, $\sum_i (X_{t_{i+1}\wedge t}-X_{t_i\wedge t})(Y_{t_{i+1}\wedge t}-Y_{t_i\wedge t})$ 依概率收敛于 $\langle X,Y\rangle_t$. 因此二次协变差有下列简单的性质: 设 M,N,M_1,M_2 是连续局部鞅.

(1) (对称性) $\langle M,N\rangle = \langle N,M\rangle$;

(2) 如果 a,b 是常数, 则 $\langle aM_1+bM_2,N\rangle = a\langle M_1,N\rangle + b\langle M_2,N\rangle$;

(3) $|\langle M,N\rangle|^2 \leqslant \langle M\rangle\langle N\rangle$.

由此可以看出, 当 M,N 是连续局部鞅时, $\langle M,N\rangle$ 是满足下面两个条件的唯一的连续有界变差过程: (i) $\langle M,N\rangle_0 = 0$; (ii) $MN-\langle M,N\rangle$ 是连续局部鞅(作为习题). 下面的定理在局部化时是非常重要的.

推论 3.4.1 设 M,N 是连续局部鞅, T 是停时, 则

$$\langle M^T,N^T\rangle = \langle M,N\rangle^T = \langle M,N^T\rangle.$$

证明 不妨假设 M,N 都是有界的. 因为 $MN-\langle M,N\rangle$ 是鞅, 故 $M^TN^T-\langle M,N\rangle^T$ 也是鞅, 即第一个等号成立. 下面我们需验证 $M^TN^T-\langle M,N^T\rangle$ 是鞅. 或者证明 $M^TN^T-MN^T$ 是鞅, 因为 $MN^T-\langle M,N^T\rangle$ 是鞅. 对任何有界停时 S, 由 Doob 停止定理,

$$\mathbb{E}(M_SN_{T\wedge S}) = \mathbb{E}(N_{T\wedge S}\mathbb{E}(M_S|\mathscr{F}_{T\wedge S})) = \mathbb{E}(N_{T\wedge S}M_{T\wedge S}),$$

即说明 $\mathbb{E}(M^TN^T-MN^T)_S=0$, 推出 $M^TN^T-MN^T$ 是鞅. 一般情况利用局部化序列容易验证. □

例 3.4.1 若 $B=(B^{(1)},\cdots,B^{(d)})$ 是 d-维 Brown 运动, 则从定理 2.6.3 的证明可以看出, $\langle B^{(i)}\rangle_t = t$. 下面我们证明当 $i\neq j$ 时, $\langle B^{(i)},B^{(j)}\rangle = 0$. 只需验证 $B^{(i)}B^{(j)}$ 是鞅就足够了. 事实上, 对 $t>s$, 由鞅性与独立性得

$$\mathbb{E}(B_t^{(i)}B_t^{(j)}-B_s^{(i)}B_s^{(j)}|\mathscr{F}_s) = \mathbb{E}((B_t^{(i)}-B_s^{(i)})(B_t^{(j)}-B_s^{(j)})|\mathscr{F}_s)$$

$$= \mathbb{E}[(B_t^{(i)} - B_s^{(i)})(B_t^{(j)} - B_s^{(j)})] = 0.$$

因此 $\langle B^{(i)}, B^{(j)} \rangle = \epsilon_{i,j} t$.

一般地, 如果 M, N 是独立的连续局部鞅, 则 $\langle M, N \rangle \equiv 0$. 这里不妨设它们是有界鞅来证明, 读者可自己用局部化方法证明一般情况. 上面的方法是没有用的, 因为一般的鞅没有独立增量性, 因此我们用定义来验证. 设 $\Delta = \{t_i\}$ 是 $[0, t]$ 的划分. 由鞅性与独立性,

$$\begin{aligned}\mathbb{E}\left(\sum_i (M_{t_i} - M_{t_{i-1}})(N_{t_i} - N_{t_{i-1}}) \right)^2 &= \mathbb{E} \sum_i (M_{t_i} - M_{t_{i-1}})^2 (N_{t_i} - N_{t_{i-1}})^2 \\ &\leqslant \mathbb{E} T_t^{\Delta}(M) \sup_i (N_{t_1} - N_{t_{i-1}})^2 \\ &\leqslant \sqrt{\mathbb{E}(T_t^{\Delta}(M))^2 \cdot \mathbb{E} \sup_i (N_{t_i} - N_{t_{i-1}})^4}.\end{aligned}$$

由定理 3.4.3 的证明知道 $\mathbb{E}(T_t^{\Delta}(M))^2$ 被一个与 Δ 无关的常数控制, 而由连续性第二项极限为零, 故 M, N 在 Δ 上的协变差的极限是零, 即 $\langle M, N \rangle = 0$.

设 M 是连续局部鞅, 写 $M_t^2 = (M_t^2 - \langle M \rangle_t) + \langle M \rangle_t$, 即 M^2 可唯一分解为一个连续局部鞅与一个连续增过程的和. 这个结果是著名的 Doob-Meyer 分解的一个特例. Doob-Meyer 分解是说一个右连续下鞅可被唯一地分解为一个右连续鞅与一个自然增过程的和, 它是现代随机分析的基石.

下面, 我们将利用连续局部鞅具有有界二次变差的事实建立随机过程相对于连续局部鞅的随机积分理论以及极其重要的 Itô 公式. 先证明一个关于积分的不等式. 设 F, G 是 $[0, \infty)$ 上递增右连续函数, α 是右连续函数, 满足对任何 $t > s \geqslant 0$,

$$|\alpha(t) - \alpha(s)|^2 \leqslant (F(t) - F(s))(G(t) - G(s)). \tag{3.4.1}$$

容易验证 α 在任何有界区间上是有限变差的且其全变差函数同样满足(3.4.1). 让我们来证明下面不等式成立.

引理 3.4.1　设 f, g 是 $[0, \infty)$ 上可测函数, 那么

$$\int_0^\infty |fg||\mathrm{d}\alpha| \leqslant \left(\int_0^\infty f^2 \mathrm{d}F \right)^{\frac{1}{2}} \left(\int_0^\infty g^2 \mathrm{d}G \right)^{\frac{1}{2}}, \tag{3.4.2}$$

其中 $|d\alpha|$ 表示 α 的全变差诱导的测度.

证明 因为全变差函数同样满足(3.4.1), 故不妨假设 α 递增且我们只需验证对任何 $t>0$ 及有界的 f,g, 有

$$\int_0^t |fg|\mathrm{d}\alpha \leqslant \left(\int_0^t f^2\mathrm{d}F\right)^{\frac{1}{2}} \left(\int_0^t g^2\mathrm{d}G\right)^{\frac{1}{2}}. \tag{3.4.3}$$

首先假设 f,g 是 $[0,t]$ 上阶梯函数, 即

$$f = H_0 1_{\{0\}} + H_1 1_{(0,t_1]} + \cdots + H_n 1_{(t_{n-1},t_n]},$$

$$g = K_0 1_{\{0\}} + K_1 1_{(0,t_1]} + \cdots + K_n 1_{(t_{n-1},t_n]},$$

其中 H_i, K_i 是常数. 由给定的条件并应用 Cauchy-Schwarz 不等式得到

$$\begin{aligned}
\int_0^t |fg|\mathrm{d}\alpha &= \sum_{i=1}^n |H_iK_i|(\alpha(t_i)-\alpha(t_{i-1}) \\
&\leqslant \sum_{i=1}^n |H_iK_i|\sqrt{F(t_i)-F(t_{i-1})}\sqrt{G(t_i)-G(t_{i-1})} \\
&\leqslant \left(\sum_{i=1}^n H_i^2(F(t_i)-F(t_{i-1}))\right)^{\frac{1}{2}} \left(\sum_{i=1}^n K_i^2(G(t_i)-G(t_{i-1}))\right)^{\frac{1}{2}} \\
&= \left(\int_0^t f^2\mathrm{d}F\right)^{\frac{1}{2}} \left(\int_0^t g^2\mathrm{d}G\right)^{\frac{1}{2}}.
\end{aligned}$$

若 μ 是 $[0,t]$ 上有限测度, 那么有界连续函数全体在 $L^2(\mu)$ 中稠. 因此阶梯函数也在 $L^2(\mu)$ 中稠. 对 $f\in L^2(\mathrm{d}F)$, $g\in L^2(\mathrm{d}G)$, 取阶梯函数列 $f_n\to f$, $g_n\to g$ (依各自的范数), 那么 $\mathrm{d}\alpha$ 几乎处处地 $f_ng_n\to fg$ (或者一个子列). 因此由 Fatou 引理,

$$\begin{aligned}
\int_0^t |fg|\mathrm{d}\alpha \leqslant \lim_n \int_0^t |f_ng_n|\mathrm{d}\alpha &\leqslant \lim_n \left(\int_0^t f_n^2\mathrm{d}F\right)^{\frac{1}{2}} \left(\int_0^t g_n^2\mathrm{d}G\right)^{\frac{1}{2}} \\
&= \left(\int_0^t f^2\mathrm{d}F\right)^{\frac{1}{2}} \left(\int_0^t g^2\mathrm{d}G\right)^{\frac{1}{2}}.
\end{aligned}$$

完成了不等式 (3.4.2) 的证明. □

设时间集为 $[0,\infty)$, $(\mathscr{F}_t)$ 是概率空间上的一个满足通常条件的流. 如果 A 是一个有界变差过程, K 是可测过程, 则我们依轨道定义

$$\left(\int_0^t K\mathrm{d}A\right)(\omega) := \int_0^t K_s(\omega)\mathrm{d}A_s(\omega),\ t \geqslant 0,\ \omega \in \Omega,$$

只要右边的积分在通常意义下存在. $t \mapsto \int_0^t K\mathrm{d}A$ 也是有界变差过程, 简记为 $K \cdot A$. 下面的不等式称为 Kunita-Watanabe 不等式.

定理 3.4.4 (Kunita-Watanabe)　设 M, N 是连续局部鞅, H, K 是可测过程, 则

$$\mathbb{E}\int_0^\infty |HK||\mathrm{d}\langle M,N\rangle| \leqslant \left(\mathbb{E}\int_0^\infty H^2\mathrm{d}\langle M\rangle\right)^{\frac{1}{2}}\left(\mathbb{E}\int_0^\infty K^2\mathrm{d}\langle N\rangle\right)^{\frac{1}{2}}. \tag{3.4.4}$$

证明　记

$$\langle M,N\rangle_s^t := \langle M,N\rangle_t - \langle M,N\rangle_s,\ t > s.$$

由二次变差的定义及初等的 Cauchy-Schwarz 不等式容易验证

$$|\langle M,N\rangle_s^t|^2 \leqslant \langle M\rangle_s^t\langle N\rangle_s^t \text{ a.s..}$$

因此由不等式(3.4.2)推出对几乎处处地, 下面的不等式成立:

$$\int_0^\infty |HK||\mathrm{d}\langle M,N\rangle| \leqslant \left(\int_0^\infty H^2\mathrm{d}\langle M\rangle\right)^{\frac{1}{2}}\left(\int_0^\infty K^2\mathrm{d}\langle N\rangle\right)^{\frac{1}{2}}. \tag{3.4.5}$$

对此不等式再应用 Cauchy-Schwarz 不等式, 我们便得到著名的 Kunita-Watanabe 不等式

$$\begin{aligned}\mathbb{E}\int_0^\infty |HK||\mathrm{d}\langle M,N\rangle| &\leqslant \mathbb{E}\left(\int_0^\infty H^2\mathrm{d}\langle M\rangle\right)^{\frac{1}{2}}\left(\int_0^\infty K^2\mathrm{d}\langle N\rangle\right)^{\frac{1}{2}}\\ &\leqslant \left(\mathbb{E}\int_0^\infty H^2\mathrm{d}\langle M\rangle\right)^{\frac{1}{2}}\left(\mathbb{E}\int_0^\infty K^2\mathrm{d}\langle N\rangle\right)^{\frac{1}{2}}.\end{aligned}$$

完成了证明.　□

现在让我们用 H^2 表示满足条件 $M_0 = 0$ 与 $\sup_{t>0}\mathbb{E}M_t^2 < \infty$ 的右连续鞅 M

全体. 设 $M \in H^2$, 则 $|M|$ 是非负下鞅, 由 Doob 不等式

$$\mathbb{E}\sup_{t>0} M_t^2 \leqslant 4\sup_{t>0}\mathbb{E}M_t^2.$$

因此 M 是一致可积的, 故容易证明 $M_\infty := \lim_{t\to\infty} M_t$ 作为 L^2-收敛极限与几乎处处收敛极限都存在, M_∞ 是平方可积的且 $M_t = \mathbb{E}(M_\infty|\mathscr{F}_t)$. 给 H^2 装备范数

$$\|M\|_{H^2} := [\mathbb{E}M_\infty^2]^{\frac{1}{2}},$$

内积为 $(M,N)_{H^2} := (M_\infty, N_\infty)_{L^2}$, 那么 $(M_t) \mapsto M_\infty$ 实际上建立了 H^2 与 $L^2(\Omega,\mathscr{F},\mathbb{P})$ 之间的等距同构. 再定义 $\langle M\rangle_\infty := \lim_{t\to\infty}\langle M\rangle_t$, 那么 $\mathbb{E}\langle M\rangle_\infty = \mathbb{E}M_\infty^2$. 用 H_c^2 表示 H^2 中的连续过程全体. 参照定理 3.4.1, 我们有下面的结论.

定理 3.4.5 H^2 是一个 Hilbert 空间且 H_c^2 是 H^2 的闭子空间.

固定 $M \in H_c^2$. 记 $L^2(M)$ 是满足

$$\|K\|_{L^2(M)}^2 := \mathbb{E}\int_0^\infty K^2 \mathrm{d}\langle M\rangle < \infty$$

的循序可测过程全体. 在这里, 如果 $[\![K - K']\!] = 0$, 我们记 $K = K'$. 对任何 $G \in \mathscr{B}[0,\infty) \times \mathscr{F}_\infty$, 定义

$$\mu_{\langle M\rangle}(G) := \mathbb{E}\int_0^\infty 1_G \mathrm{d}\langle M\rangle,$$

则 $\mu_{\langle M\rangle}$ 是 $(\mathbf{R}_+ \times \Omega, \mathscr{B}(\mathbf{R}_+) \times \mathscr{F}_\infty)$ 上的测度. 而 $L^2(M)$ 是 $L^2(\mu_{\langle M\rangle})$ 中循序可测过程组成的子空间, 容易验证 $L^2(M)$ 是闭的, 因此它也是 Hilbert 空间. 注意无论 M 是什么, 有界的左连续或右连续适应过程全体是总是 $L^2(M)$ 的子集.

定理 3.4.6 设 $M \in H_c^2$, 则对任何 $K \in L^2(M)$, 在 H_c^2 中存在唯一的元素, 记为 $K\cdot M$, 满足对任何 $N \in H_c^2$, $\langle K\cdot M, N\rangle = K\cdot\langle M,N\rangle$, 并且映射 $K \mapsto K\cdot M$ 是 $L^2(M)$ 到 H_c^2 的等距嵌入.

证明 唯一性是显然的(参考习题). 我们利用 Riesz 表示定理来证明存在性. 令

$$\phi(N) := \mathbb{E}\int_0^\infty K\mathrm{d}\langle M,N\rangle,\ N \in H_c^2.$$

应用 Kunita-Watanabe 不等式,

$$|\phi(N)| \leqslant \left(\mathbb{E}\int_0^\infty K^2 \mathrm{d}\langle M\rangle\right)^{\frac{1}{2}} \mathbb{E}\langle N\rangle_\infty = \|K\|_{L^2(M)} \cdot \|N\|_{H^2}.$$

因此 ϕ 是 Hilbert 空间 H_c^2 上的有界线性泛函. 由 Hilbert 空间上的 Riesz 表示定理, 存在唯一的元素, 记为 $K \cdot M$, 使得

$$\mathbb{E}\int_0^\infty K\mathrm{d}\langle M, N\rangle = \mathbb{E}[(K \cdot M)_\infty N_\infty], N \in H_c^2.$$

现任取有界停时 T, 在上式用 N^T 代替 N, 由推论 3.4.1 得

$$\begin{aligned}
\mathbb{E}[(K \cdot M)_T N_T] &= \mathbb{E}[(K \cdot M)_\infty N_T] \\
&= \mathbb{E}[(K \cdot M)_\infty N_\infty^T] = \mathbb{E}\int_0^\infty K\mathrm{d}\langle M, N^T\rangle \\
&= \mathbb{E}\int_0^\infty K\mathrm{d}\langle M, N\rangle^T = \mathbb{E}(K \cdot \langle M, N\rangle)_T.
\end{aligned}$$

由 Doob 有界停止定理知 $(K \cdot M)N - K \cdot \langle M, N\rangle$ 是鞅, 故 $\langle K \cdot M, N\rangle = K \cdot \langle M, N\rangle$.

另外,

$$\begin{aligned}
\|K \cdot M\|_{H^2}^2 = \mathbb{E}(K \cdot M)_\infty^2 &= \mathbb{E}\int_0^\infty K\mathrm{d}\langle M, K \cdot M\rangle \\
&= \mathbb{E}(K^2 \cdot M)_\infty M_\infty = \mathbb{E}\int_0^\infty K^2 \mathrm{d}\langle M\rangle = \|K\|_{L^2(M)}^2.
\end{aligned}$$

故映射 $K \mapsto K \cdot M$ 是等距嵌入. □

下面列出映射 $K \mapsto K \cdot M$ 的两个重要性质, 它们是定理 3.4.6 中唯一性的直接推论, 读者可自行证明.

推论 3.4.2 设 $M \in H_c^2$.

(1) 如果 $K \in L^2(M)$, $H \in L^2(K \cdot M)$, 那么 $HK \in L^2(M)$ 且 $H \cdot (K \cdot M) = (HK) \cdot M$;

(2) 如果 $K \in L^2(M)$, T 是停时, 那么

$$(K \cdot M)^T = K \cdot M^T = K^T \cdot M^T = K1_{[0,T]} \cdot M.$$

定义 3.4.1 设 $M \in H_c^2$, $K \in L^2(M)$, 则随机过程 $K \cdot M$ 称为是 K 关于 M 的

(Itô 型) 随机积分, 也写为 $\int K\mathrm{d}M$ 或 $\int_0^t K\mathrm{d}M$ 作为 t 点的值.

虽然我们已经定义了随机积分, 但还需要做两件事情:

(1) 上面随机积分中对 M 及 K 要求的条件太苛刻, 例如 Brown 运动不在 H_c^2 中, 因此我们需要拓展随机积分的定义范围;

(2) 定理 3.4.6 只是一个存在唯一性结果, 我们需要找到一种具体方法来计算随机积分.

要拓展随机积分的定义, 我们自然用局部化的方法. 设 M 是连续局部鞅, $M_0=0$. 用 $L^2_{\mathrm{loc}}(M)$ 表示满足存在趋于无穷的停时列 $\{T_n\}$, 使得

$$\mathbb{E}\int_0^{T_n} K^2\mathrm{d}\langle M\rangle<\infty,\ n\geqslant 1$$

的循序可测过程 K 的全体. 对 $K\in L^2_{\mathrm{loc}}(M)$, 我们可取趋于无穷的停时列 $\{T_n\}$, 使得对任何 n, M^{T_n} 是有界连续鞅且 $K\in L^2(M^{T_n})$. 后面总是取这样一个局部化序列.

定理 3.4.7 设 M 是连续局部鞅, $M_0=0$, $K\in L^2_{\mathrm{loc}}(M)$, 则存在唯一的连续局部鞅, 记为 $K\cdot M$, 使得 $(K\cdot M)_0=0$ 且对任何连续局部鞅 N, 有

$$\langle K\cdot M,N\rangle=K\cdot\langle M,N\rangle.$$

证明 唯一性由同定理 3.4.6 的证明类似方法推出. 为证明存在性, 取局部化序列 $\{T_n\}$, 因为 $M^{T_n}\in H^2$, $K\in L^2(M^{T_n})$, 故 $K\cdot M^{T_n}$ 有意义. 定义

$$(K\cdot M)_t:=(K\cdot M^{T_n})_t,\ t\leqslant T_n.$$

首先要证明定义无歧义. 即验证 $t\leqslant T_n$ 时, $(K\cdot M^{T_n})_t=(K\cdot M^{T_{n+1}})_t$. 由推论 3.4.2,

$$(K\cdot M^{T_n})^{T_n}=(K\cdot M^{T_n})=(K\cdot M^{T_{n+1}})^{T_n}.$$

其次由于 T_n 趋于无穷, 上式对任何 $t<\infty$ 有定义.

现在 $(K\cdot M)^{T_n}=K\cdot M^{T_n}$ 是鞅, 故 $K\cdot M$ 是局部鞅. 另外不妨设 $\{T_n\}$ 也是 N 的局部化序列,

$$\langle K\cdot M,N\rangle^{T_n}=\langle K\cdot M^{T_n},N^{T_n}\rangle$$

$$= K\cdot\langle M^{T_n}, N^{T_n}\rangle = K\cdot\langle M,N\rangle^{T_n} = (K\cdot\langle M,N\rangle)^{T_n}.$$

因此 $\langle K\cdot M, N\rangle = K\cdot\langle M,N\rangle$. □

对连续局部鞅 M, 在右边有意义时, 定义 $K\cdot M := K\cdot(M-M_0)$. 一个循序可测过程 K 称为是局部有界的, 如果存在递增趋于无穷的停时列 $\{T_n\}$ 使对任何 n, K^{T_n} 是有界过程. 局部有界的概念与 M 无关, 另外局部有界过程一定在 $L^2_{\mathrm{loc}}(M)$ 中. 左连续适应过程一定是局部有界的. 现在随机积分定义的范围已经足够大, 包括了几乎所有我们感兴趣的随机过程, 而且推论 3.4.2 中的性质仍然成立. 下面我们讨论随机积分的计算方法. 首先考虑简单的阶梯过程情况.

引理 3.4.2　设 M 是连续局部鞅, $0=t_0<t_1<\cdots<t_n$ 且 $\xi_i\in\mathscr{F}_{t_i}$, $0\leqslant i<n$, 那么

$$\left(\sum_{i=1}^n \xi_{i-1}1_{(t_{i-1},t_i]}\right)\cdot M = \sum_{i=1}^n \xi_{i-1}(M^{t_i}-M^{t_{i-1}}).$$

证明　由随机积分的线性性质, 只需证明如果 $b>a\geqslant 0$, ξ 是 $\mathscr{F}_a$ 可测的有界随机变量, 那么 $\xi 1_{(a,b]}\cdot M = \xi(M^b-M^a)$ 就足够了. 因为过程 $\xi 1_{(a,b]}$ 是左连续适应的, 故积分有意义.

我们先证明如果 (Y_t) 是鞅且当 $t\leqslant a$ 时为 0, 或简单地说 a 前等于零的鞅, 那么 ξY 是鞅. 事实上, 容易验证 ξY 是适应的, 然后任取 $t>s\geqslant 0$, 有

$$\mathbb{E}(\xi Y_t|\mathscr{F}_s) = \begin{cases} \mathbb{E}[\xi\mathbb{E}(Y_t|\mathscr{F}_a)|\mathscr{F}_s] = \mathbb{E}(\xi Y_a|\mathscr{F}_s) = 0 = Y_s, & s\leqslant a,\\ \xi\mathbb{E}(Y_t|\mathscr{F}_s) = \xi Y_s, & a<s. \end{cases}$$

因此 ξY 是连续鞅.

不妨设 M 有界, 那么 M^b-M^a 是 a 前等于零的鞅, 由此推出 $\xi(M^b-M^a)$ 是鞅. 另外对任何 $N\in H^2_c$, $(Y_tN_t-\langle Y,N\rangle_t)$ 也是 a 前等于零的鞅, 因此 $t\mapsto \xi(Y_tN_t-\langle Y,N\rangle_t)$ 也是连续鞅, 从而

$$\langle \xi Y, N\rangle = \xi\langle Y,N\rangle = \xi(\langle M,N\rangle^b - \langle M,N\rangle^a) = \xi 1_{(a,b]}\cdot\langle M,N\rangle.$$

由定理 3.4.7 的唯一性推出结论正确. □

引理中的被积过程称为是阶梯过程, 用 $\mathscr{L}_0$ 表示阶梯过程全体. 定理说明阶梯

过程的随机积分的计算方法. 下面的定理说明至少左连续适应过程的随机积分可以用类似于 Riemann 和的方法计算.

定理 3.4.8 设 M 是连续局部鞅, $\{K^{(n)}\}$ 是逐点收敛于零的局部有界过程列并且被一个局部有界过程 K 控制, 则 $(K^{(n)}\cdot M)$ 在任何有界区间上依概率一致地收敛到零. 特别地, 如果 K 是左连续适应过程, 则对任何 $t\geqslant 0$, 当 $[0,t]$ 上的分划 $\Delta=\{t_i\}$ 的长度趋于零时, 随机积分 $(K\cdot M)_t$ 是 Riemann-Stieltjes 和

$$\sum_i K_{t_i}(M_{t_{i+1}}-M_{t_i})$$

依概率收敛的极限.

证明 我们先假设 K 和 M 是有界过程. 这时 $K^{(n)}\in L^2(M)$ 并且由 Lebesgue 控制收敛定理, 它按 $L^2_{\text{loc}}(M)$ 中的度量收敛到零. 因随机积分是连续映射, 故 $K^{(n)}\cdot M$ 按 H^2_c 中的度量收敛于零, 然后 Doob 鞅不等式推出 $(K^{(n)}\cdot M)$ 在任何有界区间上依概率一致地收敛到零.

一般地, 我们只需证明对任何过程列 $X^{(n)}$, 如果存在趋于无穷的停时列 $\{T_k\}$, 使得对任何 k, $(X^{(n)})^{T_k}$ 在有限区间上依概率一致地收敛于 0, 那么 $X^{(n)}$ 也在有限区间上依概率一致地收敛于 0. 事实上, 对任何 $\epsilon>0$,

$$\begin{aligned}\mathbb{P}(\max_{s\in[0,t]}|X_s^{(n)}|>\epsilon)&\leqslant\mathbb{P}(\max_{s\in[0,t]}|X_s^{(n)}|>\epsilon,T_k\geqslant t)+\mathbb{P}(T_k\leqslant t)\\&\leqslant\mathbb{P}(\max_{s\in[0,t]}|(X^{(n)})_s^{T_k}|>\epsilon)+\mathbb{P}(T_k\leqslant t),\end{aligned}$$

因此

$$\limsup_n\mathbb{P}(\max_{s\in[0,t]}|X_s^{(n)}|>\epsilon)\leqslant\mathbb{P}(T_k\leqslant t).$$

而 $T_k\uparrow+\infty$, 故有 $\lim_n\mathbb{P}(\max_{s\in[0,t]}|X_s^{(n)}|>\epsilon)=0$.

对后一个结论, 因为 K 左连续, 故 K 是阶梯过程列

$$K^{\Delta}=K_0 1_{\{0\}}+\sum_{i=1}^{n}K_{t_{i-1}}1_{(t_{i-1},t_i]}$$

的几乎处处收敛的极限, 因此结论是定理 3.4.2 与上面结论的推论. □

此定理在实际的计算中是非常有用的.

例 3.4.2 设 $M \in H_c^2$, 我们来计算它关于自身的随机积分 $M \cdot M$. 由上面的论述, 不妨设 M 是有界的, 则 $M \in L^2_{\mathrm{loc}}(M)$. 令

$$K_t^{(n)} := \sum_{k \geqslant 0} M_{\frac{k}{2^n}} 1_{(\frac{k}{2^n}, \frac{k+1}{2^n}]}(t),\ t \geqslant 0,$$

那么对 $t \geqslant 0$, 如果我们写 $t_{n,k} := \dfrac{k}{2^n} \wedge t$, 则存在 N, 使得 $t_{n,N-1} < t_{n,N} = t$, 故

$$\begin{aligned}
\int_0^t K^{(n)} \mathrm{d}M &= \sum_{k=0}^{N-1} M_{t_{n,k}} (M_{t_{n,k+1}} - M_{t_{n,k}}) \\
&= \sum_{k=0}^{N-1} M_{t_{n,k}} M_{t_{n,k+1}} - \sum_{k=0}^{N-1} M_{t_{n,k}}^2 \\
&= \sum_{k=0}^{N-1} M_{t_{n,k}} M_{t_{n,k+1}} - \frac{1}{2}\left[\sum_{k=0}^{N-1} M_{t_{n,k}}^2 + \sum_{k=0}^{N-1} M_{t_{n,k+1}}^2 - M_t^2 + M_0^2\right] \\
&= \frac{1}{2}(M_t^2 - M_0^2) - \frac{1}{2}\sum_{k=0}^{N-1}(M_{t_{n,k+1}} - M_{t_{n,k}})^2.
\end{aligned}$$

由上一节的结果, 当 $n \to \infty$ 时,

$$\int_0^t K^{(n)} \mathrm{d}M \xrightarrow{L^2} \frac{1}{2}(M_t^2 - M_0^2 - \langle M \rangle_t),$$

因此

$$2\int_0^t M \mathrm{d}M = M_t^2 - M_0^2 - \langle M \rangle_t.$$

可以看出 $M^2 - \langle M \rangle = M_0^2 + 2M \cdot M$, 明确地给出了 M^2 的 Doob-Meyer 分解中的鞅部分. ∎

为了符号和叙述的方便, 下面我们引入半鞅的概念.

定义 3.4.2 一个连续适应过程 X 称为是一个连续半鞅, 如果它可以写为 $X = M + A$ 的形式, 其中 M 是连续局部鞅, A 是连续的有界变差过程.

因为我们总要求一个有界变差过程从零点出发, 因此一个连续半鞅 X 的如上分解是唯一的, 称为是 X 的半鞅分解, M 是 X 的鞅部分, A 是 X 的有界变差部分. 由定义和 Cauchy-Schwarz 不等式不难验证连续半鞅 X 有有界二次变差, 且

$\langle X\rangle=\langle M\rangle$. 类似定义两个连续半鞅 X, Y 的协变差过程 $\langle X, Y\rangle$, 它是协变差和的依概率收敛的极限. 设过程 K 是局部有界的, 则可自然地定义 K 关于连续半鞅 X 的随机积分:

$$K \cdot X := K \cdot M + K \cdot A,$$

其中 $X=M+A$ 是半鞅分解, K 关于 A 的积分 $K \cdot A$ 是通常意义下的按轨道的 Stieltjes 积分. 显然 $K \cdot X$ 也是一个连续半鞅, 上式恰是其半鞅分解. 当 K 是左连续过程时, 定理 3.4.8 的 Riemann-Stieltjes 和及其收敛定理仍然成立, 即 $(K \cdot X)_t$ 是 Riemann-Stieltjes 和

$$\sum_i K_{t_i}(X_{t_{i+1}\wedge t}-X_{t_i\wedge t})$$

依概率收敛的极限.

引入半鞅的一个最大的方便之处是半鞅在许多通常的运算下是封闭的, 例如两个连续鞅的乘积一般不是连续鞅, 下面的定理也称为随机积分的分部积分公式, 说明连续半鞅的乘积仍然是连续半鞅.

定理 3.4.9 设 X, Y 是两个连续半鞅, 则

$$XY=X_0Y_0+X\cdot Y+Y\cdot X+\langle X, Y\rangle.$$

证明 实际上, 上式是由其特殊形式

$$X_t^2=X_0^2+2\int_0^t X\,\mathrm{d}X+\langle X\rangle$$

极化推出. 而对 $[0,t]$ 上的有限分划 $\Delta=\{t_i\}$, 有恒等式

$$\sum_i (X_{t_{i+1}}-X_{t_i})^2=X_t^2-X_0^2-2\sum_i X_{t_i}(X_{t_{i+1}}-X_{t_i}).$$

当分划的长度趋于零时, 立即得到上面的公式. □

当 X 或 Y 是连续有界变差过程时, 我们得到通常的分部积分公式

$$XY=X_0Y_0+X\cdot Y+Y\cdot X.$$

下面我们来证明著名的 Itô 公式, 它说明连续半鞅与一个 C^2 函数复合后仍然是连

续半鞅.

定理 3.4.10(Itô) 设 $X=(X^1,\cdots,X^d)$ 是连续半鞅, $F\in C^2(\mathbf{R}^d)$, 则

$$F(X)=F(X_0)+\sum_{i=1}^d\frac{\partial F}{\partial x_i}(X)\cdot X^i+\frac{1}{2}\sum_{i,j=1}^d\frac{\partial^2 F}{\partial x_i\partial x_j}(X)\cdot\langle X^i,X^j\rangle.$$

证明 这里我们不妨只证明 $d=1$ 的情况. 设 $F\in C^2(\mathbf{R})$, X 是连续半鞅, 要证明

$$F(X)=F(X_0)+F'(X)\cdot X+\frac{1}{2}F''(X)\cdot\langle X\rangle.$$

利用定理 3.4.9, 用归纳法容易验证, 对 $n\geqslant 1$, 有

$$X^n=X_0^n+nX^{n-1}\cdot X+\frac{1}{2}n(n-1)X^{n-2}\cdot\langle X\rangle.$$

因此由随机积分的线性性, 定理当 F 是多项式时成立. 用已运用多次的局部化的方法, 我们可以设 X 是一个被常数 K 控制的有界过程. 这时存在多项式序列 $\{F_n\}$ 在闭区间 $[-K,K]$ 上依 $C^2(\mathbf{R})$ 上的范收敛于 F, 即 F_n, F_n', F_n'' 分别一致收敛于 F, F', F'', 然后由随机积分的收敛定理 3.4.8 推出 Itô 公式对 F 成立. □

Itô 公式在随机分析中的重要性是无论怎么说都不会过分的, 它可与 Newton-Leibniz 定理在微积分中的重要性相媲美.

注 1 通常我们写 $\mathrm{d}Y=H\mathrm{d}X$ 代表 $Y-Y_0=H\cdot X$, 即

$$Y_t=Y_0+\int_0^t H\mathrm{d}X,$$

则 Itô 公式可写为

$$\mathrm{d}F(X)=F'(X)\mathrm{d}X+\frac{1}{2}F''(X)\mathrm{d}\langle X\rangle,$$

故而此公式也称为链法则.

下面的定理是 Dolean-Dade 指数鞅公式的特殊形式.

定理 3.4.11 (指数鞅) 设 M 是连续局部鞅, 则对任何常数 α,

$$\mathscr{E}^\alpha(M)_t:=\exp\left\{\alpha M_t-\frac{\alpha^2}{2}\langle M\rangle_t\right\}$$

是一个连续局部鞅.

当 $\alpha=1$ 时, 写 $\mathscr{E}^1$ 为 $\mathscr{E}$.

证明 令 $F(x,y):=\exp(\alpha x-\dfrac{\alpha^2}{2}y)$, 则 $\dfrac{\partial F}{\partial y}+\dfrac{1}{2}\dfrac{\partial^2 F}{\partial x^2}=0$. 运用注 1 中高维 Itô 公式,

$$\begin{aligned}\mathscr{E}^\alpha(M)_t &= F(M_t,\langle M\rangle_t)\\ &= F(M_0,\langle M\rangle_0)+\alpha\int_0^t F(M_s,\langle M\rangle_s)\mathrm{d}M_s\\ &\quad+\int_0^t\left(\frac{\partial F}{\partial y}+\frac{1}{2}\frac{\partial^2 F}{\partial x^2}\right)(M,\langle M\rangle)\mathrm{d}\langle M\rangle.\end{aligned}$$

因此 $\mathscr{E}^\alpha(M)$ 是连续局部鞅. □

注 2 实际上, 从定理证明可看出只要 F 满足

$$\frac{\partial F}{\partial y}+\frac{1}{2}\frac{\partial^2 F}{\partial x^2}=0,$$

则 $F(M,\langle M\rangle)$ 必是一个局部鞅.

注 3 从上面的证明还可看出 $\mathscr{E}^\alpha(M)$ 实际上是随机微分方程

$$\mathrm{d}X=\alpha X\mathrm{d}M$$

的一个解. 对连续半鞅 X, 同样定义 $\mathscr{E}^\alpha(X)$, 则同样有

$$\mathscr{E}^\alpha(X)_t=\mathscr{E}^\alpha(X_0)+\alpha\int_0^t\mathscr{E}^\alpha(X)\mathrm{d}X.$$

下面 Lévy 的关于 Brown 运动的刻画定理在许多场合下的有用的.

定理 3.4.12 (Lévy) 一个从零点出发的适应的连续 d-维过程 X 是 d-维标准 Brown 运动当且仅当 X 是连续局部鞅且对任何 $1\leqslant i,j\leqslant d$, 有 $\langle X^k,X^j\rangle_t=\epsilon_{k,j}t$.

证明 设 X 是连续局部鞅且有 $\langle X^k,X^j\rangle_t=\epsilon_{k,j}t$, 则运用 Itô 公式如定理 3.4.11, 对任何 $\xi\in\mathbf{R}^d$,

$$\exp\left\{\mathrm{i}(\xi,X_t)+\frac{1}{2}|\xi|^2t\right\}$$

是连续局部鞅. 同理, 对任何 $n\geqslant 1$,

$$\exp\left\{\mathrm{i}(\xi, X_{t\wedge n}) + \frac{1}{2}|\xi|^2(t\wedge n)\right\}$$

是连续局部鞅, 因它是有界的, 故它是一个鞅. 现在对任何 $s<t$, $A\in\mathscr{F}_s$, 取 n 使得 $t\leqslant n$, 我们有

$$\mathbb{E}[1_A\exp\{\mathrm{i}(\xi, X_t - X_s)\}] = \mathbb{P}(A)\exp\left(-\frac{1}{2}|\xi|^2(t-s)\right),$$

由此推出 $X_t - X_s$ 独立于 $\mathscr{F}_s$ 并具有方差为 $t-s$ 的 Gauss 分布, 即 X 是 Brown 运动. □

Itô 公式有一个局部形式, 称为局部 Itô 公式, 也是很有用的. 设 D 是 $\mathbf{R}^d$ 的一个区域, ζ 是连续半鞅首次离开 D 的时间, 或 D^c 的首中时. 如果 $F\in C^2(D)$, 那么定理 3.4.10 中的 Itô 公式对 $t<\zeta$ 时成立. 证明也是利用局部化的方法. 设 $D_n = \{x: d(x, D^c)\geqslant \frac{1}{n}\}$, $\tau_n := \inf\{t>0:\ X_t\in D_n^c\}$, 其中 d 是 Euclid 距离, 那么 $\tau_n\uparrow\zeta$, $X^{\tau_n}\in D_n\subset D$. 因为对任何 n, $F|_{D_n}$ 总可以扩张为 $\mathbf{R}^d$ 上的 C^2 函数, 故有 Itô 公式

$$F(X^{\tau_n}) = F(X_0) + F'(X^{\tau_n})\cdot X^{\tau_n} + \frac{1}{2}F''(X^{\tau_n})\cdot\langle X^{\tau_n}\rangle.$$

即推出 Itô 公式在 $[0,\tau_n]$ 上成立, 也就是在 $[0,\zeta)$ 上成立.

例 3.4.3　让我们验证例 3.3.1 中 $h(B)$ 是 $\mathbb{P}^{x_0}$ 局部鞅的断言. 因为 $\langle B_i, B_j\rangle_t = \epsilon_{i,j}t$, 故由局部 Itô 公式, 当 $t\leqslant T_k$ 时, 有

$$\begin{aligned} h(B_t) &= h(x_0) + \int_0^t \nabla h(B)\cdot \mathrm{d}B + \frac{1}{2}\int_0^t \triangle h(B)\mathrm{d}s \\ &= h(x_0) + \int_0^t \nabla h(B)\cdot\mathrm{d}B. \end{aligned}$$

推出 $h(B)^{T_k}$ 是鞅. ▮

习　题

1. 设 M 是连续局部鞅. 证明:

   (1) $M\equiv M_0$ 当且仅当 $\langle M\rangle\equiv 0$;

   (2) $M\equiv M_0$ 当且仅当对任何连续局部鞅 N, 有 $\langle M, N\rangle\equiv 0$.

2. 设 f 是 $[0,1]$ 上的值域为至多可列集的右连续函数, 那么 f 是有界变差的当且仅当 $\sum_t |\Delta f(t)| < \infty$, 而 f 有有界二次变差当且仅当 $\sum_t |\Delta f(t)|^2 < \infty$, 其中 $\Delta f(t) = f(t) - f(t-)$.
3. 仿照二次变差的定义, 对任何 $p > 0$, 称随机过程 X 有有界 p 次变差, 如果当分划长度趋于零时, $\sum_i |X_{t_{i+1}\wedge t} - X_{t_i\wedge t}|^p$ 依概率收敛. 记 p 次变差过程为 $V^p(X)$. 设 X 是连续局部鞅, 证明: 对 $p > 2$, $V^p(X) \equiv 0$; 对 $0 < p < 2$, 则在 $\{\langle X\rangle > 0\}$ 上, $V_t^p(X) = \infty$.
4. 设 X 是一个平方可积鞅且有平稳独立增量且 $X_0 = 0$, 证明: $\langle X\rangle_t = t\cdot \mathbb{E}X_1^2$.
5. 设 M 是连续局部鞅且 $M_0 = 0$, 如果 $\langle M\rangle_\infty$ 可积, 证明: M 是 Doob 鞅.
6. 设 Z 是有界随机变量, A 是从 0 出发的有界连续增过程. 证明:

$$\mathbb{E}[ZA_\infty] = \mathbb{E}\left[\int_0^\infty \mathbb{E}(Z|\mathscr{F}_t)\mathrm{d}A_t\right].$$

7. 设 B 是标准 Brown 运动, S, T 是两个可积停时且 $S \leqslant T$. 证明: B_T 平方可积且

$$\mathbb{E}(B_T - B_S)^2 = \mathbb{E}(B_T^2 - B_S^2) = \mathbb{E}(T - S).$$

8. 设 M 是连续局部鞅. 证明: M^2 具有有界二次变差且

$$\langle M^2\rangle_t = 4\int_0^t M_s^2 \mathrm{d}\langle M\rangle_s.$$

9. 设 B 是标准 Brown 运动, X 是独立于 B 的正随机变量. 令 $M_t := B_{tX}$, $t \geqslant 0$, $(\mathscr{F}_t)$ 是一个使得 M 适应的满足通常条件的流.

   (1) 证明: M 是 $(\mathscr{F}_t)$ 局部鞅且它是一个鞅当且仅当 $\mathbb{E}X^{\frac{1}{2}} < \infty$;

   (2) 计算 $\langle M\rangle$;

   (3) 将结果推广到 $M_t = B_{A_t}$, 其中 A 是一个独立于 B 的从 0 出发的连续增过程.
10. 设 M 是有界鞅, A 是增过程, 验证:

$$\mathbb{E}\int_0^t M_s \mathrm{d}A_s = \mathbb{E}M_t A_t.$$

11. 一个下鞅 X 是类(DL)的, 如果对任何 $t \geqslant 0$, 随机变量族 $\{X_{t\wedge\tau} : \tau \in S\}$ 是一致可积的, 其中 S 是所有 $(\mathscr{F}_t)$ 停时. 证明:

(1) 鞅一定是类(DL)的;

(2) 类(DL)的局部鞅是鞅;

(3) 非负下鞅是类(DL)的.

12. 设 M 是连续局部鞅, $K \in L^2(M)$, ξ 是 $\mathscr{F}_s$ 可测随机变量. 证明: 对 $t > s \geqslant 0$,

$$\int_s^t \xi K \mathrm{d}M = \xi \int_s^t K \mathrm{d}M.$$

13. (Wald 恒等式) 设 B 是标准 Brown 运动.

(1) 如果 T 是可积停时, 证明: $\mathbb{E}B_T = 0$, $\mathbb{E}B_T^2 = \mathbb{E}T$;

(2) 设 T_b 是 b 的首中时, 证明: $\mathbb{E}T_b = \infty$.

14. 设 M, N 是连续局部鞅, $K \in L^2_{\mathrm{loc}}(M) \cap L^2_{\mathrm{loc}}(N)$, 证明: $K \cdot (aM + bN) = aK \cdot M + bK \cdot N$, $a, b \in \mathbf{R}$.

15. 设 M 是连续局部鞅. 证明: 局部有界过程在 $L^2_{\mathrm{loc}}(M)$ 中.

16. 证明: 连续适应过程是局部有界的. 举例说明右连续适应过程不一定局部有界. 问左连续适应过程是否局部有界?

17. 设 $M \in H_c^2$. 证明: 阶梯过程全体 $\mathscr{L}_0$ 在 $L^2(M)$ 中稠. 使用这个结论通过阶梯过程来定义随机积分.

18. 用归纳法验证: 对连续半鞅 X ,有

$$X_t^n = X_0^n + n\int_0^t X^{n-1} \mathrm{d}X + \frac{1}{2}n(n-1)\int_0^t X^{n-2} \mathrm{d}\langle X \rangle.$$

19. 设 X, Y 是两连续半鞅, 在什么条件下, $\mathscr{E}(X+Y) = \mathscr{E}(X)\mathscr{E}(Y)$?

20. 设 X, Y 是两个连续半鞅, 定义

$$X \circ Y := X \boldsymbol{\cdot} Y + \frac{1}{2}\langle X, Y \rangle.$$

它称为 X 关于 Y 的 Stratonovich 积分, 也记为 $\int_0^t X \circ \mathrm{d}Y$.

(1) $\int_0^t X \circ \mathrm{d}Y$ 是和式

$$\sum_{i=1}^n \frac{X_{t_i} + X_{t_{i-1}}}{2}(Y_{t_i} - Y_{t_{i-1}})$$

当分划 $0 = t_0 < t_1, \cdots < t_n = t$ 的长度趋于零时依概率收敛的极限.

(2) 证明: $XY - X_0Y_0 = X \circ Y + Y \circ X$.

21. 设 M 是连续局部鞅使得测度 $\mathrm{d}\langle M\rangle_t$ 几乎处处与 Lebesgue 测度等价, 证明: 存在循序可测过程 V 和 Brown 运动 B, 使得

$$M_t = M_0 + \int_0^t V_s \mathrm{d}B_s.$$

22. 设 B 是 1-维 Brown 运动, K 局部可积. 那么 $X = K \cdot B$ 是连续局部鞅. 用 K 写出时间变换 τ_t, 使得 X_τ 是 Brown 运动. 验证 X 是 τ-连续的.
23. 设 M 是实连续局部鞅, 证明: M 在 $\{\sup_t M_t < \infty\}$ 上几乎处处收敛.
24. 单连通区域 $D \subset \mathbf{R}^d$ 上的函数 f 是调和的当且仅当对任何 $x \in D$, $f(B^\tau)$ 是 $\mathbb{P}^x$-鞅, 其中 B 是标准 Brown 运动, τ 是 D^c 的首中时.

§3.5 Girsanov 公式与鞅表示

这一节中, 我们将继续讨论半鞅与随机积分的一些性质. 设 $(\Omega, \mathscr{F}, \mathbb{P})$ 是一个概率空间, $(\mathscr{F}_t)$ 是其上的满足通常条件的流, 令 $\mathscr{F}_\infty := \sigma(\bigcup_{t\geqslant 0} \mathscr{F}_t)$.

对增过程 A, 我们定义 A 的右连续逆

$$\tau_t := \inf\{s : A_s > t\}, \ t \geqslant 0,$$

则 (τ_t) 是增的右连续的一族 $(\mathscr{F}_t)$-停时. 当 A 连续时, $A_{\tau_t} = t$. 令 $\widehat{\mathscr{F}_t} := \mathscr{F}_{\tau_t}$, $t \geqslant 0$, 那么 A_t 是 $(\widehat{\mathscr{F}_t})$-停时. 一般地, 一个增的右连续的停时族 $(\tau_t : t \geqslant 0)$ 称为是一个时间变换. 实际上, 实际变换与增过程是一一对应的. 为了简单起见, 下面我们总假设 $\lim_{t\to\infty} A_t = \infty$, 或者说 $\zeta := \inf\{t : \tau_t = \infty\} = \infty$. 这个假设并非本质需要的, 只是如果没有这个假设, 下面的结论都只能在 $t < \zeta$ 上成立.

设 X 是 $(\mathscr{F}_t)$ 循序可测的, 则时间变换后的过程 $\widehat{X}_t := X_{\tau_t}$ 是 $(\widehat{\mathscr{F}_t})$-适应的. 称 X 是 τ-连续的, 如果 X 在 $[\tau_-, \tau]$ 上是常数. 显然如果 X 是连续的且 τ-连续, 则 $\widehat{X}$ 是连续的. 下面引理的证明留作习题.

引理 3.5.1 设 M 是连续局部鞅, $S \leqslant T$ 是停时, 则 M 在 $[S,T]$ 是常数当且仅当 $\langle M\rangle$ 在 $[S,T]$ 上是常数.

引理 3.5.2　如果 M 是 τ-连续的, 连续 $(\mathscr{F}_t)$ 局部鞅, 那么 $\widehat{M}$ 是连续的 $(\widehat{\mathscr{F}_t})$ 局部鞅, 且 $\langle\widehat{M}\rangle = \widehat{\langle M\rangle}$.

证明　首先设 M 是有界的. 我们只需验证 $\widehat{M}$ 是 $(\widehat{\mathscr{F}_t})$ 鞅. 对任何 $n \geqslant 1$, $s < t$, 由 Doob 有界停止定理,

$$\mathbb{E}(M_{\tau_t \wedge n} | \mathscr{F}_{\tau_s}) = M_{\tau_s \wedge n}.$$

再由有界收敛定理推出 $\mathbb{E}(M_{\tau_t}|\mathscr{F}_{\tau_s}) = M_{\tau_s}$, 即证明之. 同样, 因 $M^2 - \langle M\rangle$ 是有界连续鞅且 τ-连续, 故 $\widehat{M}^2 - \widehat{\langle M\rangle}$ 是连续鞅. 因此由唯一性推出 $\langle\widehat{M}\rangle = \widehat{\langle M\rangle}$.

一般地, 取停时 T 使得 M^T 是有界的, 那么 $\widehat{T} := \inf\{t : \tau_t \geqslant T\}$ 是 $(\widehat{\mathscr{F}_t})$ 停时(习题)且 $\tau_{\widehat{T}-} \leqslant T \leqslant \tau_{\widehat{T}}$. 因为 M 是 τ-连续的, 故 M 在 $[T, \tau_{\widehat{T}}]$ 上是常数, 推出下面等式的最后一个等号

$$\widehat{M}_t^{\widehat{T}} = \widehat{M}_{\widehat{T}\wedge t} = M_{\tau_{\widehat{T}\wedge t}} = M_{T\wedge \tau_t}.$$

因此 $\widehat{M^T} = \widehat{M}^{\widehat{T}}$, 即知道后者是 $(\widehat{\mathscr{F}_t})$ 有界连续鞅且

$$\langle\widehat{M^T}\rangle = \widehat{\langle M\rangle^T} = \widehat{\langle M\rangle}^{\widehat{T}}.$$

如果 $\{T_n\}$ 是局部化停时列, 由定义及前面的假设推出 $\widehat{T}_n \uparrow \infty$. 由此也推出了定理的结论. □

下面定理说明任何连续局部鞅经一个时间变换后成为 Brown 运动.

定理 3.5.1 (Dambis, Dubins-Schwarz)　如果 M 是连续局部鞅, $M_0 = 0$ 且设 $\langle M\rangle_\infty = \infty$. 如果 (τ_t) 是 $\langle M\rangle$ 的右连续逆, 那么 $\widehat{M}$ 是标准 Brown 运动且 $M = B_{\langle M\rangle}$, 其中 $B := \widehat{M}$.

证明　由引理 3.5.1, 因为 $\langle M\rangle$ 是 τ-连续的, 故 M 也是. 由引理 3.5.2, $B_t := M_{\tau_t}$ 是连续局部鞅且 $\langle B\rangle = \langle M\rangle_{\tau_t}$. 因 $\langle M\rangle$ 连续, 故 $\langle M\rangle_{\tau_t} = t$ 即 $\langle B\rangle_t = t$, 因此 B 是标准 Brown 运动.

另外 $B_{\langle M\rangle} = M_{\tau_{\langle M\rangle}}$, 因为 $\tau_{\langle M\rangle_t -} \leqslant t \leqslant \tau_{\langle M\rangle_t}$, 但 M 是 τ-连续的, 因此对任何 $t \geqslant 0$,

$$M_{\tau_{\langle M\rangle_t}} = M_t,$$

即 $M = B_{\langle M\rangle}$. □

Itô 公式是对 C^2 中函数成立的, 但在一维情况下, 它可以通过到凸函数. 设 f 是 $\mathbf{R}$ 上一个凸函数, 自然 f 是连续的, 左导数 f'_- 存在且是递增的, 它诱导一个 $\mathbf{R}$ 上的测度 μ_f. 这时, Itô 公式是怎么样的呢?

定理 3.5.2　若 X 是连续半鞅, f 是凸函数, 则存在连续增过程 A^f, 使得对任何 t 几乎处处

$$f(X_t) = f(X_0) + \int_0^t f'_-(X)\mathrm{d}X + \frac{1}{2}A_t^f.$$

证明　取 $j \in C_c^\infty((-\infty, 0))$ 满足 $\int j\mathrm{d}x = 1$. 置

$$f_n(x) := \int_{-\infty}^0 f\left(x + \frac{y}{n}\right) j(y)\mathrm{d}y,$$

那么 $f_n \in C^\infty$ 凸, $f_n \longrightarrow f$ 且 $f'_n \uparrow f'_-$. 由 Itô 公式

$$f_n(X_t) = f_n(X_0) + \int_0^t f'_n(X)\mathrm{d}X + \frac{1}{2}A_t^{f_n},$$

其中 $A_t^{f_n} = \int_0^t f''_n(X)\mathrm{d}\langle X\rangle$. 上式左边 a.s. 收敛于 $f(X_t)$. 由局部化, 不妨设 X 与 $f'_-(X)$ 都是有界的, 因此由定理 3.4.8, $\int_0^t f'_n(X)\mathrm{d}X$ 在任何紧区间上一致地依概率收敛于 $\int_0^t f'_-(X)$. 故 $\int_0^t f'_n(X)\mathrm{d}X$ (至少有一个子列) a.s. 收敛于 $\int_0^t f'_-(X)\mathrm{d}X$. 令

$$A_t^f := f(X_t) - f(X_0) - \int_0^t f'_-(X)\mathrm{d}X,$$

那么 A^f 是连续的且 $A_t^{f_n}$ a.s. 收敛于 A_t^f. 因为 A^{f_n} 是增过程, 故 A^f 也是增过程. □

用 $\mathrm{sgn}(\cdot)$ 表示 $|x|$ 的左导数函数.

推论 3.5.1 (Tanaka)　对任何 $a \in \mathbf{R}$, 存在连续增过程 L^a, 使得

$$|X_t - a| = |X_0 - a| + \int_0^t \mathrm{sgn}(X - a)\mathrm{d}X + L_t^a.$$

另外, 测度 $\mathrm{d}L_t^a$ 支持在集合 $\{t : X_t = a\}$ 上.

证明　只需证明第二个断言. 应用 Itô 公式和上面定理于 $|X - a|$, 得

$$
\begin{aligned}
(X_t-a)^2 &= (X_0-a)^2+2\int_0^t |X-a|d|X-a|+\langle|X-a|\rangle \\
&= (X_0-a)^2+2\int_0^t |X-a|\operatorname{sgn}(X-a)\mathrm{d}X+2\int_0^t |X-a|\mathrm{d}L_t^a+\langle X\rangle \\
&= (X_0-a)^2+2\int_0^t (X-a)\mathrm{d}X+\langle X\rangle+2\int_0^t |X-a|\mathrm{d}L_t^a \\
&= (X_t-a)^2+2\int_0^t |X-a|\mathrm{d}L_t^a,
\end{aligned}
$$

因此 $\int|X-a|\mathrm{d}L^a=0$. 推出结论. □

过程 L^a 通常称为 X 在 a 点的局部时, 也记为 $L^a(X)$. 下面的 Itô-Tanaka 公式说明了凸函数的二阶导数作为测度的意义,参考文献 [29], §4.5. 设 f 是凸函数,则

$$
f(X_t)=f(X_0)+\int_0^t f'_-(X)\mathrm{d}X+\frac{1}{2}\int_{\mathbf{R}} L_t^a\mu_f(\mathrm{d}a),
$$

其中 μ_f 是 f 在分布意义下的二阶导数.

下面讨论半鞅在测度变换下的变化. 设 $\mathbb{Q}$ 是 $(\Omega,\mathscr{F})$ 的另一个概率测度, 称 $\mathbb{Q}$ 关于 $\mathbb{P}$ 局部绝对连续, 如果对任何 $t\in\mathsf{T}$, $\mathbb{Q}|_{\mathscr{F}_t}$ 关于 $\mathbb{P}|_{\mathscr{F}_t}$ 绝对连续; 称 $\mathbb{Q}$ 关于 $\mathbb{P}$ 绝对连续, 如果 $\mathbb{Q}|_{\mathscr{F}_\infty}$ 关于 $\mathbb{P}|_{\mathscr{F}_\infty}$ 绝对连续. 一个过程的鞅性与所讲的测度有关, 因此我们用 $\mathbb{P}$- (或 $\mathbb{Q}$-) 表示相对于测度 $\mathbb{P}$ (或 $\mathbb{Q}$).

现在设 $\mathbb{Q}$ 关于 $\mathbb{P}$ 局部绝对连续, 记 $Z=(Z_t)$ 为局部密度, 即在 $\mathscr{F}_t$ 上, $\mathbb{Q}=Z_t\cdot\mathbb{P}$, $t\in\mathsf{T}$.

定理 3.5.3　设 $\mathbb{Q}$ 关于 $\mathbb{P}$ 局部绝对连续, 且 $Z=(Z_t)$ 为局部密度. 那么

(1) Z 是 $\mathbb{P}$-鞅, 且它是一致可积当且仅当 $\mathbb{Q}$ 关于 $\mathbb{P}$ 绝对连续;

(2) 一个适应过程 X 是 $\mathbb{Q}$-鞅当且仅当 $X\cdot Z$ 是 $\mathbb{P}$-鞅.

证明　显然 Z 非负, 且 $\mathbb{E}Z_t=\mathbb{Q}1=1$. 对任何 $t>s$, $A\in\mathscr{F}_s\subset\mathscr{F}_t$,由局部密度的定义, 有 $\mathbb{E}(Z_t;A)=\mathbb{Q}(A)=\mathbb{E}(Z_s;A)$, 因此 Z 是 $\mathbb{P}$-鞅. 如果 X 是 $\mathbb{Q}$-鞅, 那么 XZ 是 $\mathbb{P}$-可积的且

$$
\mathbb{E}(X_tZ_t;A)=\mathbb{Q}(X_t;A)=\mathbb{Q}(X_s;A)=\mathbb{E}(X_sZ_s;A),
$$

因此 XZ 是 $\mathbb{P}$-鞅. 此方法是可逆的, 故 (2) 得证.

下面我们证明 (1) 的第二部分. 由鞅的正则化定理, 我们可设 Z 是右连续的. 若 Z 一致可积, 则 Z 是一个可闭的鞅, 即存在 $Z_\infty \in F_\infty$, 使得 $Z_t \xrightarrow{L^1} Z_\infty$ 且 $Z_t = \mathbb{E}(Z_\infty|\mathscr{F}_t)$. 对任何 t, $A \in \mathscr{F}_t$, 有

$$\mathbb{Q}(A) = \mathbb{E}(Z_t; A) = \mathbb{E}(Z_\infty; A),$$

推出上式对 $\mathscr{F}_\infty$ 成立, 即 $\mathbb{Q}$ 关于 $\mathbb{P}$ 绝对连续. 反之容易推出 Z 是可闭鞅, 因此是一致可积的. □

下面我们设 $\mathbb{Q}$ 关于 $\mathbb{P}$ 局部绝对连续且不失一般性地设局部密度 Z 是右连续的.

定理 3.5.4　对任何停时 τ, 有 $\mathbb{Q} = Z_\tau \cdot \mathbb{P}$ 在 $\mathscr{F}_\tau \cap \{\tau < \infty\}$ 上成立. 另外, 一个适应右连续过程 X 是 $\mathbb{Q}$-局部鞅, 如果 XZ 是 $\mathbb{P}$-局部鞅.

证明　对任何 $t \in \mathsf{T}$, $A \in \mathscr{F}_\tau$, 有 $A \cap \{\tau < t\} \in \mathscr{F}_{\tau\wedge t}$, 且由 Doob 停止定理, $\mathbb{E}(Z_t|\mathscr{F}_{\tau\wedge t}) = Z_{\tau\wedge t}$. 因此

$$\begin{aligned}\mathbb{Q}(A \cap \{\tau < t\}) &= \mathbb{E}(Z_t; A \cap \{\tau < t\})\\ &= \mathbb{E}(Z_{\tau\wedge t}; A \cap \{\tau < t\})\\ &= \mathbb{E}(Z_\tau; A \cap \{\tau < t\}),\end{aligned}$$

让 $t \to +\infty$, 得 $\mathbb{Q}(A \cap \{\tau < +\infty\}) = \mathbb{E}(Z_\tau; A \cap \{\tau < +\infty\})$.

对第二个断言, 只需证 X^τ 是 $\mathbb{Q}$-鞅当且仅当 $(XZ)^\tau$ 是 $\mathbb{P}$-鞅. 令 $\mathscr{F}_t^\tau := \mathscr{F}_{\tau\wedge t}$, 那么相对于流 $(\mathscr{F}_t^\tau)$, $\mathbb{Q}$ 关于 $\mathbb{P}$ 局部绝对连续且局部密度是 Z^τ. 因此 X^τ 是 $(\mathscr{F}_t)$ 的 $\mathbb{Q}$-鞅当且仅当 X^τ 是 $(\mathscr{F}_t^\tau)$ 的 $\mathbb{Q}$-鞅, 由定理 3.5.3, 这当且仅当 $X^\tau Z^\tau$ 是 $(\mathscr{F}_t^\tau)$ 的 $\mathbb{P}$-鞅, 当且仅当是 $(\mathscr{F}_t)$ 的 $\mathbb{P}$-鞅. 如果 XZ 是 $\mathbb{P}$ 局部鞅, 存在停时列 $\tau_n \uparrow +\infty$ a.s. $\mathbb{P}$, 使得 $(XZ)^{\tau_n}$ 是 $\mathbb{P}$ 鞅, 那么 X^{τ_n} 是 $\mathbb{Q}$-鞅. 而且容易推出 $\tau_n \uparrow +\infty$ a.s. $\mathbb{Q}$. 因此 X 是 $\mathbb{Q}$-局部鞅. □

下面的引理说明局部密度函数关于测度 $\mathbb{Q}$ 是严格正的.

引理 3.5.3　局部密度 Z 是严格正的 $\mathbb{Q}$-a.s., 即除一个 $\mathbb{Q}$ 零测集外, $t \mapsto Z_t$ 是严格正的.

证明　不妨设 $\mathbb{Q}$ 关于 $\mathbb{P}$ 绝对连续, 否则可以考虑用固定时间 t 停止. 令 $T := \inf\{t > 0: Z_t = 0\}$. 只需验证 $\mathbb{Q}(T < \infty) = 0$. 事实上, 因 Z 是非负 $\mathbb{P}$-鞅, 由定理

3.3.4, 在 $\{t > T\}$ 上, $Z_t = 0$ a.s.($\mathbb{P}$), 即推出在 $\{T < \infty\}$ 上, $Z_\infty = 0$ a.s.($\mathbb{P}$), 那么 $\mathbb{Q}(T < \infty) = \mathbb{P}(Z_\infty; \{T < \infty\}) = 0$. □

下面是著名的 Girsanov 公式.

定理 3.5.5 (Girsanov) 设 $\mathbb{Q}$ 关于 $\mathbb{P}$ 局部绝对连续且局部密度是连续的. 那么对任何连续 $\mathbb{P}$-局部鞅 M, 过程 $\widetilde{M} := M - \dfrac{1}{Z} \cdot \langle M, Z\rangle$ 是一个连续 $\mathbb{Q}$-局部鞅.

证明 如果 Z^{-1} 是有界的, 那么 $\widetilde{M}$ 是一个连续的 $\mathbb{P}$-半鞅, 显然 $\langle \widetilde{M}, Z\rangle = \langle M, Z\rangle$, 由分部积分公式得

$$\begin{aligned}\widetilde{M}_t Z_t - \widetilde{M}_0 Z_0 &= \int_0^t \widetilde{M} \mathrm{d}Z + \int_0^t Z d\widetilde{M} + \langle \widetilde{M}, Z\rangle_t \\ &= \int_0^t \widetilde{M} \mathrm{d}Z + \int_0^t Z \mathrm{d}M - \langle M, Z\rangle_t + \langle \widetilde{M}, Z\rangle_t \\ &= \int_0^t \widetilde{M} \mathrm{d}Z + \int_0^t Z \mathrm{d}M,\end{aligned}$$

即 $\widetilde{M} Z$ 是一个 $\mathbb{P}$-局部鞅, 由定理 3.5.4, $\widetilde{M}$ 是 $\mathbb{Q}$-局部鞅.

对一般的 Z, 令 $\tau_n := \inf\{t > 0 : Z < \dfrac{1}{n}\}$, 则 τ_n 是停时且由上面的引理, $\tau_n \uparrow +\infty$ a.s. $\mathbb{Q}$. 而 $\dfrac{1}{Z^{\tau_n}}$ 有界, 故推出 $\widetilde{M}^{\tau_n}$ 是一个 $\mathbb{Q}$-局部鞅, 因此 $\widetilde{M}$ 也是一个 $\mathbb{Q}$-局部鞅. □

下面我们将反过来证明给定一个 $\mathbb{P}$-鞅 Z, 则有一个关于 $\mathbb{P}$ 局部绝对连续的概率 $\mathbb{Q}$ 以 Z 作为局部密度. 在实际中这种情况更经常地出现.

定理 3.5.6 设 E 是 Polish 空间, $X = (X_t)$ 是 $(\Omega, \mathscr{F}, \mathbb{P})$ 上以 $(E, \mathscr{E})$ 为状态空间的随机过程, $\mathscr{F}_t^0 := \sigma(\{X_s : s \leqslant t\})$, $\mathscr{F}_\infty^0 := \sigma(\{X_s : s \in \mathsf{T}\})$. 若 $Z = (Z_t)$ 是 $(\mathscr{F}_t^0)$ 非负鞅且 $\mathbb{E}Z_0 = 1$, 那么存在 $(\Omega, \mathscr{F}_\infty)$ 上的概率测度 $\mathbb{Q}$, 使得对任何 $t \geqslant 0$, 在 $\mathscr{F}_t^0$ 上 $\mathbb{Q} = Z_t \cdot \mathbb{P}$.

证明 对任何 $t \in \mathsf{T}$, $A \in \mathscr{F}_t^0$, 定义 $\mathbb{Q}(A) := \mathbb{E}(Z_t; A)$. Z 的鞅性保证 $\mathbb{Q}$ 在 Ω 的代数 $\mathscr{A} := \bigcup_{t\in\mathsf{T}} \mathscr{F}_t^0$ 上的定义无歧义且 $\mathbb{E}Z_0 = 1$ 保证 $\mathbb{Q}(\Omega) = 1$. 由测度扩张定理, 只需验证 $\mathbb{Q}$ 在 $\mathscr{A}$ 上有上连续性就足够了. 因为以上的 $\mathbb{Q}$ 的定义实际上确定了 E 上的一个相容的有限维分布族, 由 Kolmogorov 定理, 在典则空间 $(E^\mathsf{T}, \mathscr{E}^\mathsf{T})$ 上存在概

率测度 $\overline{\mathbb{P}}$, 使得典则过程的有限维分布族与 X 在 $\mathbb{Q}$ 下的有限维分布族一致, 因此推出 $\mathbb{Q}$ 也有上连续性. □

在上节中, 我们看到如果 M 是一个局部鞅且 $M_0=0$, 那么其指数鞅

$$\mathscr{E}(M)=\exp\left(M-\frac{1}{2}\langle M\rangle\right)$$

也是一个局部鞅且 $\mathbb{E}\mathscr{E}(M)_0=1$. 实际上逆命题也成立, 即若 Z 是严格正的连续局部鞅, 则存在连续局部鞅 M, 使得 $Z=\mathscr{E}(M)$. 但一般地, Z 仅是一个上鞅, 不是鞅. 可是为了定理 3.5.6 成立, $\mathscr{E}(M)$ 必须是一个真正的鞅, 而不仅仅是局部鞅. 由 Wald 恒等式(见习题)可得到一个充分条件: (Novikov) 若 M 是连续局部鞅, $M_0=0$ 且对任何 $t\geqslant 0$, $\mathbb{E}\mathrm{e}^{\frac{\langle M\rangle_t}{2}}<\infty$, 则 $\mathbb{E}\mathscr{E}(M)_t=1$, $t\geqslant 0$, 即 $\mathscr{E}(M)$ 是一个鞅.

例 3.5.1 (Cameron-Martin) 设 N 是连续局部鞅, $Z:=\mathscr{E}(N)$ 是鞅, $\mathbb{Q}$ 是关于 $\mathbb{P}$ 局部绝对连续的概率测度使得 Z 是局部密度. 再设 M 是 $\mathbb{P}$-连续局部鞅, 那么 $\widetilde{M}_t=M_t-Z_t^{-1}\cdot\langle M,Z\rangle=M_t-\langle M,N\rangle_t$ 是 $\mathbb{Q}$-连续局部鞅. 特别地, 设 B 是 $(\Omega,\mathscr{F},\mathbb{P})$ 上标准 Brown 运动, 如果 $H=(H_s)$ 是有界循序可测过程, 那么

$$Z_t:=\exp\left(\int_0^t H_s\mathrm{d}B_s-\frac{1}{2}\int_0^t H_s^2\mathrm{d}s\right)$$

是鞅, 令 $\mathbb{Q}=Z\cdot\mathbb{P}$, 那么过程 $\widetilde{B}_t:=B_t-\int_0^t H_s\mathrm{d}s$ 是 $\mathbb{Q}$-连续局部鞅而且 $\langle\widetilde{B}\rangle_t=\langle B\rangle_t=t$. 因此由定理 3.4.12, 过程 $\widetilde{B}$ 是关于测度 $\mathbb{Q}$ 的标准 Brown 运动. 或者说, 原 Brown 运动在新的测度下是一个半鞅, 其分解为 $B_t=\widetilde{B}_t+\int_0^t H_s\mathrm{d}s$. ▮

下面我们讨论相对于 Brown 运动流的鞅及其表示. 设 $B=(B_t)$ 是概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 上的标准 d-维 Brown 运动, $(\mathscr{F}_t)$ 是 B 的自然流经过完备化后得到的流, 当然它满足通常条件.

定理 3.5.7 对任何 $\xi\in L^2(\Omega,\mathscr{F}_\infty,\mathbb{P})$, 存在唯一的适应过程 $H\in L^2(B)$, 使得

$$\xi=\mathbb{E}\xi+\int_0^\infty H_s\mathrm{d}B_s.$$

证明 我们只对 $d=1$ 证明. 不妨设 $\mathbb{E}\xi=0$, 否则用 $\xi-\mathbb{E}\xi$ 代替. 定义

$$\mathscr{H}:=\{\xi\in L^2(\Omega,\mathscr{F}_\infty,\mathbb{P}):\mathbb{E}\xi=0\},$$

$\mathscr{H}'$ 是 $\mathscr{H}$ 中可以如定理中那样表示的 ξ 全体, 那么对 $\xi \in \mathscr{H}'$, 有

$$\mathbb{E}\xi^2 = \mathbb{E}\int_0^\infty H_s^2 \mathrm{d}s.$$

唯一性是显然的且由此可验证 $\mathscr{H}'$ 是 $\mathscr{H}$ 的闭子空间. 现在要验证如果 $\xi \perp \mathscr{H}'$, 那么 $\xi = 0$ a.s..

任取 $f \in L^2(\mathbf{R})$, 用 $\mathscr{E}(f)$ 表示 $if \cdot B$ 的指数鞅, 那么它满足方程

$$\mathscr{E}_t(f) - 1 = i\int_0^t \mathscr{E}_s(f)f(s)\mathrm{d}B_s.$$

因为右边属于 $\mathscr{H}'$, 故

$$\mathbb{E}[\xi \cdot \mathscr{E}_\infty(f)] = \mathbb{E}(\xi(\mathscr{E}_\infty(f) - 1)) = 0.$$

取 f 是阶梯函数

$$f = \sum_{j=1}^n y_j 1_{(t_{j-1}, t_j]},$$

其中 $0 = t_0 < t_1 < \cdots < t_n$, $y_1, \cdots, y_n \in \mathbf{R}$, 那么

$$\mathbb{E}\left(\xi \cdot \exp\left\{i\sum_{j=1}^n y_j(B_{t_j} - B_{t_{j-1}})\right\}\right) = 0$$

对任何 $\{t_j\}$ 和 $\{y_j\}$ 成立. 由特征函数的唯一性推出, 对任何 $H \in \mathscr{B}(\mathbf{R}^n)$, 有

$$\mathbb{E}(\xi; (B_{t_1} - B_{t_0}, \cdots, B_{t_n} - B_{t_{n-1}}) \in H) = 0,$$

那么自然也有

$$\mathbb{E}(\xi; (B_{t_1}, \cdots, B_{t_n}) \in H) = 0.$$

然后由单调收敛定理和 Dynkin 定理, 对任何 $A \in \mathscr{F}_\infty$, 有 $\mathbb{E}(\xi; A) = 0$, 故 $\xi = 0$ a.s. 完成证明. □

定理 3.5.8 (鞅表示定理) 任何 $(\mathscr{F}_t)$ 局部鞅 M 的一个修正可以表示为

$$M_t = C + \int_0^t H_s \mathrm{d}B_s,$$

其中 C 是常数, H 是局部地在 $L^2(B)$ 中的适应过程. 特别地, 任何 $(\mathscr{F}_t)$ 局部鞅有连续的修正.

证明 先设 M 是 L^2 有界的, 那么 $M_\infty \in L^2(\Omega, \mathscr{F}_\infty, \mathbb{P})$. 由上面的定理存在 $H \in L^2(B)$, 使得

$$M_\infty = \mathbb{E}M_\infty + \int_0^\infty H_s \mathrm{d}B_s.$$

因此

$$M_t = \mathbb{E}(M_\infty|\mathscr{F}_t) = C + \int_0^t H_s \mathrm{d}B_s, \text{ a.s.},$$

右边是连续的, 说明 M 有一个连续的修正.

然后证明当 M 是一致可积鞅时也有连续修正. 因为这时候, 存在 L^2 有界的鞅 $M^{(n)}$ 使得 $M_\infty^{(n)} \to M_\infty$. 由 Doob 不等式和 Borel-Cantelli 引理推出 $M^{(n)}$ 有一个子列几乎处处地一致收敛于 M.

最后设 M 是 $(\mathscr{F}_t)$ 局部鞅. 那么它可以局部化为一致可积鞅, 因此有一个连续修正. 再将这个修正局部化为连续有界鞅推出它可以如上表示. □

习 题

1. 设 M 是连续局部鞅, $S \leqslant T$ 是停时.

   (1) 如果 $K_t := \xi 1_{(S,T]}$, 其中 ξ 是有界 $\mathscr{F}_S$ 可测的, 则 $K \cdot M = \xi(M^T - M^S)$;

   (2) 证明: $\langle M \rangle_S = \langle M \rangle_T$ 当且仅当 M 在 $[S,T]$ 上是常值 (引理 3.5.1).

2. 设 M 是连续局部鞅, $M_0 = 0$, L 是 0 点的局部时. 证明:

   (1) $\inf\{t : L_t > 0\} = \inf\{t : \langle M \rangle_t > 0\}$ a.s.;

   (2) 当 $0 < \alpha < 1$, $M \not\equiv 0$ 时, $|M|^\alpha$ 不是半鞅.

3. 证明: 在 Tanaka 公式的假设下,

$$(X_t - a)^+ = (X_0 - a)^+ + \int_0^t 1_{\{X>a\}} \mathrm{d}X + \frac{1}{2}L_t^a,$$

$$(X_t - a)^- = (X_0 - a)^- - \int_0^t 1_{\{X \leqslant a\}} \mathrm{d}X + \frac{1}{2}L_t^a.$$

4. 设 $\mathbb{Q}$ 关于 $\mathbb{P}$ 局部绝对连续, 则 $\mathbb{P}$-连续半鞅 X 是 $\mathbb{Q}$-连续半鞅且 $\langle X\rangle_{\mathbb{Q}} = \langle X\rangle_{\mathbb{P}}$. 另外有界循序可测过程 H 关于 X 的 $\mathbb{Q}$-随机积分与 $\mathbb{P}$-随机积分一致.
5. (Wald 恒等式) 设 B 是标准 Brown 运动. 如果停时 T 满足 $\mathbb{E}e^{\frac{T}{2}} < \infty$, 证明: $\mathbb{E}\exp\left(B_T - \dfrac{T}{2}\right) = 1$. $\Bigg($提示: 令 $Z_t := \exp\left(B_t - \dfrac{t}{2}\right)$. 用 Girsanov 公式证明对停时 $T_b = \inf\{t \geqslant 0:\ B_t = t - b\}$, $b > 0$ 成立. 然后验证 Z^{T_b} 是 Doob 鞅.$\Bigg)$

§3.6 随机微分方程

在这一节的后面, 我们将简单地介绍下列形式的随机微分方程的一些结果

$$\mathrm{d}X_t = b(t, X)\mathrm{d}t + \sigma(t, X)\mathrm{d}B_t, \tag{3.6.1}$$

其中 B 是一个 r-维标准 Brown 运动, X 是未知的连续 d-维过程. 实际上是下列随机积分方程成立

$$X_t - X_0 = \int_0^t b(s, X)\mathrm{d}s + \int_0^t \sigma(s, X)\mathrm{d}B_s. \tag{3.6.2}$$

我们先对系数 (b, σ) 作个说明.

设 W^d 是 $[0, +\infty)$ 到 $\mathbf{R}^d$ 的连续映射全体, 装备紧一致收敛的拓扑, 用 $\mathscr{B}(W^d)$ 表示 W^d 上 Borel σ- 代数, $\mathscr{B}_t(W^d)$ 表示由 $w \mapsto w(s)$, $s \in [0, t]$ 生成的子 σ-代数, $\mathbf{R}^d \otimes \mathbf{R}^r$ 表示 $d \times r$ 矩阵全体, $\mathscr{A}^{d,r}$ 表示满足下列条件的可测映射 $\alpha : [0, \infty) \times W^d \longrightarrow \mathbf{R}^d \otimes \mathbf{R}^r$ 全体: 对任何 $t \geqslant 0$, $\alpha(t, \cdot)$ 是 $(W^d, \mathscr{B}_t(W^d))$ 到 $\mathbf{R}^d \otimes \mathbf{R}^r$ 可测的. 自然地, 我们要求 $\sigma \in \mathscr{A}^{d,r}$, $b \in \mathscr{A}^{d,1}$. 这时, X 是 $(\mathscr{F}_t)$-适应的蕴含着 $\sigma(t, X)$ 是 $(\mathscr{F}_t)$-循序可测的. 当 $\sigma(t, X) = \sigma(t, X_t)$, $b(t, X) = b(t, X_t)$ 时(这时右边的 $\sigma : \mathbf{R}_+ \times \mathbf{R}^d \longrightarrow \mathbf{R}^d \otimes \mathbf{R}^r, b : \mathbf{R}_+ \times \mathbf{R}^d \longrightarrow \mathbf{R}^d \otimes \mathbf{R}$), 我们称 (3.6.1) 是 Markov 型的; 而当 $\sigma(t, X) = \sigma(X_t)$, $b(t, X) = b(X_t)$ 时(这时右边的 $\sigma : \mathbf{R}^d \longrightarrow \mathbf{R}^d \otimes \mathbf{R}^r, b : \mathbf{R}^d \longrightarrow \mathbf{R}^d \otimes \mathbf{R}$), 称 (3.6.1) 是 Itô 型的或扩散型的.

另外, 我们需要解释所谓 (3.6.1) 的解的意义. 给定 σ, b 如上, 随机微分方程 (3.6.1)(也称为方程 (σ, b)) 的解是带有流 $(\mathscr{F}_t)$ 的概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的适应过程,

对 (X, B) 满足:

(1) B 是 $\mathbf{R}^r$ 上标准 $(\mathscr{F}_t)$-Brown 运动;

(2) (3.6.1) 成立.

解的唯一性有两种不同的解释. 我们说 (3.6.1) 的解有轨道唯一性是指两个有相同概率空间, 流, 适应 Brown 运动及几乎处处相等的初值的解是不可区分的, 即若在同一个带流的概率空间上的两个解 (X, B) 和 (X', B') 且若 $B = B'$, $X_0 = X'_0$, 则 $X = X'$. 另外我们说解有分布唯一性是指两个具相同初始分布的解 X 和 X' 是等价的, 即有相同的有限维分布族. 有时我们也区分强解与弱解, 解 (X, B) 称为强解, 如果 X 关于 B 生成的且完备化的流 $(\mathscr{F}_t^B)$ 适应. 非强解的解称为弱解. 实际上, 强解是指概率空间及其上的 Brown 运动是给定时, 满足 (3.6.1) 的过程 X.

定理 3.6.1 如果 (3.6.1) 有轨道唯一性, 那么

(1) 分布唯一性也成立;

(2) 每个解都是强解, 实际上, 存在 $F: W^r \longrightarrow W^d$, 对任何 $t \geqslant 0$, F 是 $(W^r, \mathscr{B}_t(W^r))$ 到 $(W^d, \mathscr{B}_t(W^d))$ 的可测映射, 使得 $X = F(B)$.

下面为了简单, 我们讨论 Itô 型的方程, 一些结果在一般的情况也成立, 其证明也无需很大的修改. 读者可参考文献 [18] 或文献 [21].

定义 $a := \sigma\sigma^T$ 及椭圆型算子

$$Af(x) := \frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}(x) D_i D_j f(x) + \sum_{i=1}^{d} b_i(x) D_i f(x),\ \ f \in C_c^\infty(\mathbf{R}^d),\ \ x \in \mathbf{R}^d,$$

其中 $D_i := \dfrac{\partial}{\partial x_i}$. 一个连续过程 X 是鞅问题 (a, b) 的解是指对任何 $f \in C_c^\infty(\mathbf{R}^d)$, 过程

$$M_t(f) := f(X_t) - f(X_0) - \int_0^t Af(X_s)\mathrm{d}s$$

是局部鞅. (3.6.1) 的解存在性和上述鞅问题的解的存在性是一样的.

定理 3.6.2 (Stroock, Varadhan) 设 σ, b 可测, 那么 (3.6.1) 的解存在当且仅当上述鞅问题的解存在.

证明 先设 (3.6.1) 的解存在为 (X, B), 则由 Itô 公式得

$$df(X_t) = \sum_{i=1}^{d} D_i f(X_t) \mathrm{d}X_t^i + \frac{1}{2} \sum_{i,j=1}^{d} D_i D_j f(X_t) \mathrm{d}\langle X^i, X^j \rangle_t$$
$$= Df(X_t)\sigma(X_t)\mathrm{d}B_t + Af(X_t)\mathrm{d}t.$$

因此 X 是上述鞅问题的解.

反之, 如果 X 是上述鞅问题的解, 取函数 $f(x) = x_i$, 用局部化的方法容易验证

$$M_t := X_t - X_0 - \int_0^t b(X_s)\mathrm{d}s, \ t \geqslant 0$$

是 d-维连续局部鞅. 再取函数 $f(x) = x_i x_j$, 及局部化方法可以证明

$$M_t^{i,j} := X_t^i X_t^j - X_0^i X_0^j - \int_0^t a_{i,j}(X)\mathrm{d}s - \int_0^t X_s^i b_j(X)\mathrm{d}s - \int_0^t X_s^j b_i(X)\mathrm{d}s$$

是连续局部鞅. 由分部积分公式

$$M^{i,j} = X^i \cdot M^j + X^j \cdot M^i + \langle M^i, M^j \rangle - \int_0^{\cdot} a_{i,j}(X)\mathrm{d}s,$$

因此推出

$$\langle M^i, M^j \rangle_t = \int_0^t a_{i,j}(X_s)\mathrm{d}s.$$

由表示定理(参考文献 [21] 的定理 16.12 或文献 [29] 的定理 V.3.9), 存在 r-维标准 Brown 运动 B, 使得

$$M_t = \int_0^t \sigma(X_s)\mathrm{d}B_s.$$

因此 (X, B) 是随机微分方程 (3.6.1) 的解. □

例 3.6.1 (Langevin 方程)　考虑方程

$$\mathrm{d}X_t = \mathrm{d}B_t - X_t \mathrm{d}t. \tag{3.6.3}$$

分部积分, 得

$$\mathrm{d}\mathrm{e}^t X_t = \mathrm{e}^t \mathrm{d}X_t + X_t \mathrm{e}^t \mathrm{d}t = \mathrm{e}^t \mathrm{d}B_t.$$

因此 $\mathrm{e}^t X_t - X_0 = \int_0^t \mathrm{e}^s \mathrm{d}B_s$ 或 $X_t = \mathrm{e}^{-t} X_0 + \mathrm{e}^{-t} \int_0^t \mathrm{e}^s \mathrm{d}B_s$, 是著名的 Ornstein-

Uhlenbeck 过程. 因此方程强解存在且有轨道唯一性.

例 3.6.2 (几何Brown运动) 考虑方程

$$\mathrm{d}X_t = bX_t\mathrm{d}t + \sigma X_t\mathrm{d}B_t,$$

其中 b, σ 为常数. 由 Itô 公式

$$\begin{aligned}\mathrm{d}(\ln X_t) &= \frac{1}{X_t}\mathrm{d}X_t - \frac{1}{2X_t^2}\mathrm{d}\langle X\rangle_t \\ &= b\mathrm{d}t + \sigma\mathrm{d}B_t - \frac{1}{2X_t^2}\sigma^2 X_t^2\mathrm{d}t \\ &= \left(b - \frac{\sigma^2}{2}\right)\mathrm{d}t + \sigma\mathrm{d}B_t.\end{aligned}$$

因此

$$\ln X_t - \ln X_0 = \left(b - \frac{\sigma^2}{2}\right)t + \sigma B_t,$$

即解得

$$X_t = X_0 \exp\left(\left(b - \frac{\sigma^2}{2}\right)t + \sigma B_t\right).$$

例 3.6.3 (圆周Brown运动) 设 B 是 Brown 运动, $X = (\cos B, \sin B)$. 由 Itô 公式,

$$\mathrm{d}X_1(t) = -\sin B_t\mathrm{d}B_t - \frac{1}{2}\cos B_t\mathrm{d}t;$$

$$\mathrm{d}X_2(t) = \cos B_t\mathrm{d}B_t - \frac{1}{2}\sin B_t\mathrm{d}t.$$

因此 $X = (X_1, X_2)$ 满足下列随机微分方程组

$$\mathrm{d}X_1 = -\frac{1}{2}X_1\mathrm{d}t - X_2\mathrm{d}B,$$

$$\mathrm{d}X_2 = -\frac{1}{2}X_2\mathrm{d}t + X_1\mathrm{d}B.$$

例 3.6.4 设 $\sigma(x) = \mathrm{sgn}(x)$, $x \in \mathbf{R}$, 考虑方程

$$\mathrm{d}X_t = \sigma(X_t)\mathrm{d}B_t.$$

任取概率空间及其上的标准 Brown 运动 $\tilde{B}$, $\mathscr{F}_0$ 可测随机变量 ξ, 令

$$X = \xi + \tilde{B},\ B_t = \int_0^t \sigma(X)\mathrm{d}\tilde{B},$$

那么 B 是连续鞅且 $\langle B\rangle_t = \int_0^t \sigma^2(X)\mathrm{d}s = t$, 因此 B 也是标准 Brown 运动且

$$\mathrm{d}X = \mathrm{d}\tilde{B} = \sigma^2(X)\mathrm{d}\tilde{B} = \sigma(X)\mathrm{d}B,$$

即方程的解存在(但不是强解, 参见 [29] 的 第 9 章 §1), 而且有分布唯一性, 因为任何解都和 Brown 运动有相同的有限维分布, 但没有轨道唯一性, 因为如果 X 是初值为零的解, 那么 $-X$ 也是. ▮

像通常微分方程一样, 我们一般几乎不可能写出随机微分方程的显式解, 但仿照常微分方程的 Picard 迭代方法, 我们也可以证明当 σ, b 满足局部 Lipschitz 条件时, 方程的解存在且有轨道唯一性.

定理 3.6.3　设存在常数 C, 使得

(1) $|b(x)| + |\sigma(x)| \leqslant C(1+|x|),\ x \in \mathbf{R}^d$;

(2) $|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leqslant C|x-y|,\ x, y \in \mathbf{R}^d$,

ξ 是独立于给定 Brown 运动 B 的平方可积随机变量, 则 (3.6.1) 有解 X 满足 $X_0 = \xi$ 且有轨道唯一性.

注　因为 (3.6.1) 当 $\sigma = 0$ 是通常的常微分方程, 因此两个条件即使从常微分方程的意义上也是必要的. 条件 (1) 保证方程的解不在有限时间内 blow up; 条件 (2) 保证解的轨道唯一性. □

证明　先证明唯一性, 设 X, Y 是具相同初值 ξ 的两个解, 那么

$$\begin{aligned}\mathbb{E}|X_t - Y_t|^2 &= \mathbb{E}|\int_0^t b(X_s) - b(Y_s)\mathrm{d}s + \int_0^t \sigma(X_s) - \sigma(Y_s)\mathrm{d}B_s|^2 \\ &\leqslant C\int_0^t \mathbb{E}|X_s - Y_s|^2\mathrm{d}s.\end{aligned}$$

令 $v(s) := \mathbb{E}|X_s - Y_s|^2$, 则 $v(t) \leqslant C\int_0^t v(s)\mathrm{d}s$, $v(0) = 0$, 由 Gronwall 不等式(见习题)推出 $v \equiv 0$, 即 $X_t = Y_t$ a.s., 因此 X, Y 不可区别.

为证明存在性, 迭代定义

$$Y_t^{(0)} := \xi,\ Y_t^{(k+1)} := \xi + \int_0^t b(Y_s^{(k)})\mathrm{d}s + \sigma(Y_s^{(k)})\mathrm{d}B_s. \tag{3.6.4}$$

对 $t \in [0,T]$, 由条件 (1) 与 (2) 计算得

$$\mathbb{E}|Y_t^{(k+1)} - Y_t^{(k)}|^2 \leqslant (1+T)3C^2 \cdot \int_0^t \mathbb{E}|Y_s^{(k)} - Y_s^{(k-1)}|^2\mathrm{d}s,\ k \geq 1,$$

$$\mathbb{E}|Y_t^{(1)} - Y_t^{(0)}|^2 \leqslant 2C^2t^2(1+|\xi|^2) + 2C^2t(1+\mathbb{E}|\xi|^2) \leqslant C_1 t.$$

利用归纳法验证

$$\mathbb{E}|Y_t^{(k+1)} - Y_t^{(k)}| \leqslant \frac{C_2^{k+1}t^{k+1}}{(k+1)!},\ k \geqslant 0, t \in [0,T],$$

其中 C_1, C_2 仅依赖 C, T 及 $\mathbb{E}|\xi|^2$. 由鞅不等式及条件 (2),

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\in[0,T]}|Y_t^{(k+1)} - Y_t^{(k)}| > \frac{1}{2^k}\right)\\
\leqslant{}& \mathbb{P}\left(\int_0^T |b(Y_s^{(k)}) - b(Y_s^{(k-1)})|\mathrm{d}s + \sup_{t\in[0,T]}\left|\int_0^t \sigma(Y_s^{(k)}) - \sigma(Y_s^{(k-1)})\mathrm{d}B_s\right| > \frac{1}{2^k}\right)\\
\leqslant{}& \mathbb{P}\left(\int_0^T |b(Y_s^{(k)}) - b(Y_s^{(k-1)})|\mathrm{d}s > \frac{1}{2^{k+1}}\right)\\
&+ \mathbb{P}\left(\sup_{t\in[0,T]}\left|\int_0^t \sigma(Y_s^{(k)}) - \sigma(Y_s^{(k-1)})\mathrm{d}B_s\right| > \frac{1}{2^{k+1}}\right)\\
\leqslant{}& 2^{2k+2}T\int_0^T \mathbb{E}|b(Y_s^{(k)}) - b(Y_s^{(k-1)})|^2\mathrm{d}s + 2^{2k+2}\int_0^T \mathbb{E}|\sigma(Y_s^{(k)}) - \sigma(Y_s^{(k-1)})|^2\mathrm{d}s\\
\leqslant{}& 2^{2k+2}C^2(T+1)\int_0^T \frac{C_2^k t^k}{k!}\mathrm{d}s \leqslant \frac{(4C_2T)^{k+1}}{(k+1)!}.
\end{aligned}$$

由 Borel-Cantelli 引理, 对几乎所有的轨道, 当 k 充分大时

$$\sup_{t\in[0,T]}\left|Y_t^{(k+1)} - Y_t^{(k)}\right| \leqslant \frac{1}{2^k},$$

因此 $Y_t^{(n)} = Y_t^{(0)} + \sum\limits_{k=1}^{n}(Y_t^{(k)} - Y_t^{(k-1)})$ 在 $[0,T]$ 上一致收敛, 记其极限为 X. 因 $Y^{(n)}$ 是连续适应过程, 故 X 也是.

下面证明 X 满足 (3.6.1). 在 (3.6.3) 中让 $k \to \infty$, 左边的极限为 X_t. 由 Fatou 引理,

$$\mathbb{E}\int_0^T |X_t - Y_t^{(k)}|^2 \mathrm{d}t \leqslant \liminf_m \mathbb{E}\int_0^T |Y_t^{(m)} - Y_t^{(k)}|^2 \mathrm{d}t \longrightarrow 0,$$

那么 $\mathbb{E}\int_0^t |\sigma(X_s) - \sigma(Y_s^{(k)})|^2 \mathrm{d}s \longrightarrow 0$. 由随机积分的性质推出

$$\int_0^t \sigma(Y_s^{(k)})\mathrm{d}B_s \xrightarrow{L^2} \int_0^t \sigma(X_s)\mathrm{d}B_s.$$

再由 Hölder 不等式推出

$$\int_0^t b(Y_s^{(k)})\mathrm{d}s \xrightarrow{L^2} \int_0^t b(X_s)\mathrm{d}s.$$

这样我们推出 X 满足 (3.6.1) 且 $X_0 = \xi$. □

对于 Markov 型的方程有类似的结果, 只要 b, σ 满足相应的条件: 对任何 T, 存在常数 C, 使得

(1) $|b(t,x)| + |\sigma(t,x)| \leqslant C(1+|x|),\ x \in \mathbf{R}^d,\ t \in [0,T]$;

(2) $|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leqslant C|x-y|,\ x, y \in \mathbf{R}^d,\ t \in [0,T]$.

其证明没有实质的不同. 但在 1 维时, 上述条件可以本质地减弱, 如下面的定理. 证明参考文献 [18], [22].

定理 3.6.4 (Yamada-Watandbe) 考虑一维 Markov 型方程, 即 $d = r = 1$, 满足对所有的 $t \geqslant 0$, $x, y \in \mathbf{R}$.

(1) $|b(t,x) - b(t,y)| \leqslant C \cdot |x-y|$;

(2) $|\sigma(t,x) - \sigma(t,y)| \leqslant h(|x-y|)$,

其中 $h: \mathbf{R}_+ \to \mathbf{R}_+$ 严格增, $h(0) = 0$ 且 $\int_{0+} \dfrac{\mathrm{d}x}{h^2(x)} = \infty$. 那么方程有轨道唯一性.

上述方程通常称为由 Brown 运动驱动的, 关于更一般的方程, 读者可阅读相应的参考书.

习 题

1. 解方程 $\mathrm{d}X_t = r\mathrm{d}t + \alpha X_t \mathrm{d}B_t$. (提示: 利用因子 $\exp(-\alpha B_t + \frac{1}{2}\alpha^2 t)$.)

2. (线性方程) 设 U, V 是连续半鞅, $Z := \exp(V - V_0 - \frac{1}{2}\langle V\rangle)$, 证明: 方程 $\mathrm{d}X = \mathrm{d}U + X\mathrm{d}V$ 有唯一解 $X = Z(X_0 + Z^{-1}\cdot(U - \langle U, V\rangle))$.

3. 设 σ, b, c 是适当维数的函数, 证明: 方程 $\mathrm{d}X = b(X)\mathrm{d}t + \sigma(X)\mathrm{d}B$ 与方程 $\mathrm{d}X = (b(X) + \sigma(X)c(X))\mathrm{d}t + \sigma(X)\mathrm{d}B$ 的解同时存在或不存在. (用 Girsanov 定理.)

4. (Gronwall 不等式) 设 $\mathbf{R}$ 上连续函数 g 满足:

$$0 \leqslant g(t) \leqslant \alpha(t) + \beta\int_0^t g(s)\mathrm{d}s,\ t \in [0, T],$$

其中 $\beta \geqslant 0$, $\alpha : [0, T] \to \mathbf{R}$ 可积. 证明:

$$g(t) \leqslant \alpha(t) + \beta\int_0^t \alpha(s)\mathrm{e}^{\beta(t-s)}\mathrm{d}s,\ t \in [0, T].$$

第四章 Markov 过程基础

在第二章中我们引入了转移半群及 Markov 过程, 在这一章中我们将更进一步讨论 Markov 过程, 引入 Markov 过程的两个非常重要的性质:右连续性与强 Markov 性(通常被称为是 Meyer 右假设). 它们是现代 Markov 过程与概率位势理论研究的基础. 然后我们证明 Feller 过程满足这些性质, 另外还将更详细地讨论 Brown 运动与 Lévy 过程的一些性质.

§4.1 右 Markov 过程

在 §2.3 中, 我们是从转移函数的角度用 Kolomogorov 相容定理来构造 Markov 过程, 它是满足 Markov 性的一族有序随机变量. 但是这样的 Markov 过程对于我们的研究来说太过于一般, 我们希望 Markov 过程有更好的性质使得我们可以讨论概率位势理论. 在这一节中, 我们将引入右 Markov 过程的概念, 粗略地说, 右 Markov 过程就是样本轨道右连续且满足强 Markov 性的时间齐次 Markov 过程, 真正要说清楚它的定义却颇费周折, 有很多可测性细节需要说明. 直观地把 Markov 过程想象成为在确定的一个随机规则下运动的粒子, 它的可能到达的点全体称为状态空间, 用 E 表示. 它的轨迹称为是样本轨道, 是 $[0,\infty)$ 到 E 的映射, 样本轨道的全体是样本空间, 用 Ω 表示. 粒子可以在某个时刻消失, 这时我们说它进入了坟墓 Δ. 对任何 $x\in E$, 概率 $\mathbb{P}^x$ 描述从 x 出发的样本轨道的行为, 随机规则就是这样的一族概率. Markov 性是指如果粒子在时刻 t 到达状态 x, 那么它忘记了过去的行为, 从 x 点依概率 $\mathbb{P}^x$ 重新开始运动.

设 $(E,\mathscr{E})$ 是一个可测空间. 对于 E 上的任何概率测度 μ, 记 $\mathscr{E}^\mu$ 是 E 关于 μ 的完备化. 所有如此完备化的交

$$\mathscr{E}^* := \bigcap_{\mu} \mathscr{E}^\mu$$

称为是 E 上普遍可测集的 σ- 代数, 也称为是 $\mathscr{E}$ 的普遍完备化. 任何 $(E,\mathscr{E}^*)$ 上的概率可唯一延拓为 $(E,\mathscr{E}^*)$ 上的概率. 不难证明, f 是 $\mathscr{E}^*$-可测当且仅当对任何 E 上的概率 μ, 存在 $\mathscr{E}$ 可测的 f_1, f_2, 满足 $f_1 \leqslant f \leqslant f_2$ 且 $\mu(f_1 \neq f_2) = 0$. (见习题.) 对于任何拓扑空间, 我们认为其有一个自然的 Borel σ- 代数与普遍可测的 σ- 代数. 现在设 E 是一个 Radon 空间, 即它可以同胚于一个紧度量空间 $\hat{E}$ 的一个普遍可测子集, 用 $C_u(E)$ 表示 E 上有界一致连续函数全体. 设 $\mathscr{E}$ 是 E 上的 Borel σ- 代数, 后面我们将看到这样的一般性假设的必要性. 任何 $\mathscr{E}$ 上的概率测度可唯一地延拓到 $\mathscr{E}^*$ 上. 我们说 (P_t) 是 E 上的一个转移半群是指它是 $(E,\mathscr{E}^*)$ 上转移半群. 如果需要强调, $(E,\mathscr{E})$ 上转移半群称为是 Borel 转移半群. 显然 Borel 转移半群可唯一地延拓成为 $(E,\mathscr{E}^*)$ 上转移半群. 我们不把自己限制于 Borel 转移半群的理由是即使在最简单的过程变换下, 一个 Borel 转移半群都不一定仍然是 Borel 转移半群. 下面我们将叙述 Meyer 的两个右假设.

任意固定点 $\Delta \notin E$. 如果 E 是 LCCB (局部紧具可数基的 Hausdorff)空间, Δ 取为 E 的 Alexander 紧化点. $E_\Delta := E \cup \{\Delta\}$. 一个映射 $\omega : [0,\infty] \longrightarrow E_\Delta$ 称为是一个样本轨道, 如果

(1) $\omega(\infty) = \Delta$;

(2) 存在 $\zeta(\omega) \in [0,\infty]$ 使得当 $t \geqslant \zeta(\omega)$ 时, $\omega(t) = \Delta$; 当 $t < \zeta(\omega)$ 时, $\omega(t) \in E$;

(3) 当 $t < \zeta(\omega)$ 时, ω 是右连续的.

对样本轨道 ω, $\zeta(\omega)$ 称为是样本轨道 ω 的生命时(或者死亡时). 对任何 t, $X_t(\omega)$ 表示 ω 在 t 处的值, 即 $X_t(\omega) = \omega(t)$. 对任何 s, 定义

$$\theta_s\omega(t) = \omega(t+s),\ t \geqslant 0.$$

显然 $\theta_s\omega \in \Omega$. $\theta_s\omega$ 称为是 s 处的推移轨道.

用 Ω 表示以上定义的样本轨道全体, 那么对任何 $s \geqslant 0$, $\theta_s(\Omega) \subset \Omega$, $(\theta_s : s \in \mathbf{R}_+)$ 称为是 Ω 上的推移算子族. 定义 $\mathscr{F}^{0*} := \sigma(X_t : t \geqslant 0)$, 其中 X_t 看成为 Ω 到 $(E,\mathscr{E}^*)$ 的映射. 类似地定义 $\mathscr{F}_t^{0*} := \sigma(X_s : 0 \leqslant s \leqslant t)$ 过程 (X_t) 的自然流. 用 $\mathscr{F}^0$ 和 $\mathscr{F}_t^0$ 表示将 X_t 看成为 Ω 到 $(E,\mathscr{E})$ 的映射生成的自然流以示区别. 显然 $\mathscr{F}^0 \subset \mathscr{F}^{0*}$, $\mathscr{F}_t^0 \subset \mathscr{F}_t^{0*}$. 将恒等于 Δ 的轨道记为 $[\,\Delta\,]$, 我们总设所有样本空间包含 $[\,\Delta\,]$. 且约定一个 E 上的函数在 Δ 上赋值为零, 一个 Ω 上随机变量在 $[\,\Delta\,]$ 上的值为零.

考虑 $X(t,\omega)=X_t(\omega)$ 作为 $\mathbf{R}_+\times\Omega$ 到 E_Δ 的映射. 对任何 $n\geqslant 1$, 定义

$$X_n(t,\omega)=X\left(\frac{j+1}{2^n},\omega\right),\ \frac{j}{2^n}<t\leqslant\frac{j+1}{2^n},$$

那么 $X_n(t,\omega)$ 作为 $(\mathbf{R}_+\times\Omega,\mathscr{B}(\mathbf{R}_+)\times\mathscr{F}^{0*})$ 到 $(E,\mathscr{E}^*)$ 是可测的且 X_n 点点收敛于 X. 因此 $(t,\omega)\mapsto X(t,\omega)$ 也是可测的.

定义 4.1.1 设 $(P_t)_{t\geqslant 0}$ 是 E 上的一个转移半群. 称它满足右假设 1 (HD1), 如果对 $(E,\mathscr{E})$ 上任何概率 μ, 存在 $(\Omega,\mathscr{F}^{0*})$ 上的概率 $\mathbb{P}^\mu$, 满足:

(1) 对任何 $A\in\mathscr{E}^*$, 有 $\mathbb{P}^\mu(X_0\in A)=\mu(A)$;

(2) 对任何 $t,s\geqslant 0$, $f\in b\mathscr{E}^*$, 有

$$\mathbb{E}^\mu(f(X_{t+s})|\mathscr{F}_s^{0*})=P_tf(X_s).$$

对任何 $x\in E$, 简单地记 $\mathbb{P}^{\epsilon_x}$ 为 $\mathbb{P}^x$, 称为是 x 点出发的概率. 性质 (1) 说明 μ 是过程的初始分布, (2) 说明过程是以 (P_t) 为转移半群的时齐 Markov 过程. 由单调类定理容易验证 (2) 等价于对任何有界或非负 $\mathscr{F}^{0*}$-可测随机变量 Y, 有

$$\mathbb{E}^x(Y\circ\theta_s|\mathscr{F}_s^{0*})=\mathbb{E}^{X_s}(Y).$$

过程 X 取值在 E_Δ, 但是因为它一旦进入 Δ 就不再活动, 所以它形象地被称为坟墓点, 并且我们通常说过程的状态空间是 E. 过程 X 称为保守的, 如果对任何 $x\in E$, $\mathbb{P}^x(\zeta<\infty)=0$. 这时过程在有限时间内不会进入坟墓. 保守性等价于对任何 $t>0$, $x\in E$ 有 $\mathbb{P}^x(X_t\in E)=1$. 取上面的典则型样本空间来定义 (HD1) 并非必需, 只是为了叙述的方便. 回忆 §2.3 中右连续简单 Markov 过程的定义, 容易看出, 如果 (HD1) 成立, 那么

$$X=(\Omega,\mathscr{F}^{0*},\mathscr{F}_t^{0*},X_t,\theta_t,\mathbb{P}^x)$$

是以 $(E_\Delta,\mathscr{E}_\Delta^*)$ 为状态空间的右连续简单 Markov 过程, 不难验证它满足对任何 $t\geqslant 0$ 有 $\mathbb{P}^\Delta(X_t\in E)=0$. 反之, 设 X 是这样的一个右连续简单 Markov 过程, 因为 $x\mapsto\mathbb{P}^x(A)$ 对任何 $A\in\mathscr{F}^{0*}$ 是 $\mathscr{E}^*$ 可测的, 故对 $(E,\mathscr{E}^*)$ 上任何概率 μ 定义 $\mathbb{P}^\mu:=\int_E\mathbb{P}^x\mu(\mathrm{d}x)$, 然后将测度族 $\{\mathbb{P}^\mu\}$ 通过典则映射成为样本轨道空间上的概率, 满足 (HD1). 任何右连续简单 Markov 过程都等价于典则型样本空间上的一个满足 (HD1) 的 Markov 过程. 因此一个转移半群满足(HD1)当且仅当它有一个右连续实

现, 或者说半群 (P_t) 满足 (HD1) 当且仅当它是一个右连续简单 Markov 过程的转移半群.

如果 f 是有界连续函数, 则 $P_tf(x)=\mathbb{E}^x f(X_t)$ 是 t 的右连续函数. 因为 X 在概率 $\mathbb{P}^x$ 下的有限维分布族由 (P_t) 唯一决定, 即概率族 $\{\mathbb{P}^x\}$ 由 (P_t) 唯一决定. 要注意的是我们不知道什么条件能够刻画出转移半群由这样的一个 Markov 过程来产生. 下一节中讲述的 Feller 性质是一个充分条件.

引理 4.1.1 对任何 $A\in\mathscr{E}^*$, 则对任何 $\mathbf{R}_+$ 和 E 上的概率测度 λ 和 μ, 映射 $(t,x)\mapsto P_t(x,A)$ 是关于 $(\mathscr{B}(\mathbf{R}_+)\times\mathscr{E}^*)$ 经测度 $\lambda\times\mu$ 完备化后的 σ- 代数可测的.

证明 如果 f 是有界连续的, 那么 $P_tf(x)=\mathbb{E}^x f(X_t)$ 关于 t 右连续. 故这时 $(t,x)\mapsto P_tf(x)$ 关于 $\mathscr{B}(\mathbf{R}_+)\times\mathscr{E}^*$ 可测. 然后由单调类定理可以证明当 f 是有界 $\mathscr{E}$-可测函数时, 它有同样的可测性. 现在对以上的 λ 和 μ, 定义

$$\nu(\cdot):=\int\lambda(\mathrm{d}t)\int\mu(\mathrm{d}x)P_t(x,\cdot)$$

是 E 上的有限测度. 对 $\mathscr{E}^*$ 可测的 f, 存在 $\mathscr{E}$ 可测的 f_1,f_2, 满足 $f_1\leqslant f\leqslant f_2$ 且 $\nu(f_2\neq f_1)=0$, 那么 $P_tf_1(x)\leqslant P_tf(x)\leqslant P_tf_2(x)$, 两端的函数是 $\lambda\times\mu$ 几乎处处相等且都是 $\mathscr{B}(\mathbf{R}_+)\times\mathscr{E}^*$ 可测的, 因此映射 $(t,x)\mapsto P_t(x,A)$ 是关于 $(\mathscr{B}(\mathbf{R}_+)\times\mathscr{E}^*)$ 经测度 $\lambda\times\mu$ 完备化后的 σ- 代数可测的. □

类似地, 可以证明对任何 $A\in\mathscr{E}^*$, $t\mapsto P_t(x,A)$ 对任何 x 是 Lebesgue 可测的. 因此对 $\alpha>0$, 我们可以定义

$$U^\alpha f(x):=\int_0^\infty \mathrm{e}^{-\alpha t}P_tf(x)\mathrm{d}t,\ f\in\mathscr{E}^*_+,$$

称为位势算子、豫解算子或 Green 算子. 因为 $(t,\omega)\mapsto f(X_t(\omega))$ 是可测的, 故由 Fubini 定理,

$$U^\alpha f(x)=\mathbb{E}^x\int_0^\infty \mathrm{e}^{-\alpha t}f(X_t)\mathrm{d}t.$$

容易验证:

(1) $U^\alpha 1\leqslant\dfrac{1}{\alpha}$;

(2) 若 f 有界且在 x 点连续, 那么

$$\lim_{\alpha\to\infty}\alpha U^\alpha f(x)=f(x).$$

到现在为止流 $(\mathscr{F}_t^{0*})$ 是自然流, 与概率 $\{\mathbb{P}^x\}$ 无关. 因为它不满足通常条件, 故 Borel 集的首中时都不一定是 $(\mathscr{F}_t^{0*})$ 停时. 为了方便让我们用概率来强化 $(\mathscr{F}_t^{0*})$, 这样不同的概率会得到不同的强化, 流与具体的概率有关了.

强化是与完备化类似但不同的概念. 设 $(\Omega,\mathscr{F},\mathbb{P})$ 是一个完备概率空间, $\mathscr{A}$ 是 $\mathscr{F}$ 的一个子 σ- 代数, $\mathscr{A}$ (关于 $(\Omega,\mathscr{F},\mathbb{P})$ 的)强化是指 $\tilde{\mathscr{A}} := \sigma(\mathscr{A}\cup\mathscr{N})$, 其中 $\mathscr{N}$ 是 $\mathbb{P}$ 零概率集全体. 强化后的 σ- 代数一般比完备化后的 σ- 代数大. (参考附录.)

设 μ 是 E 上概率测度, 用 $\mathscr{F}^\mu$ 表示 $\mathscr{F}^{0*}$ 关于 $\mathbb{P}^\mu$ 的完备化, $\mathscr{N}^\mu$ 表示 $\mathscr{F}^\mu$ 中的 $\mathbb{P}^\mu$-零概率集全体. 令

$$\mathscr{F} := \bigcap_\mu \mathscr{F}^\mu,\ \ \mathscr{N} := \bigcap_\mu \mathscr{N}^\mu.$$

用 $\mathscr{F}_t^\mu$ 表示 $\mathscr{F}_t^{0*}$ 关于 $(\Omega,\mathscr{F}^\mu,\mathbb{P}^\mu)$ 的强化, 即 $\mathscr{F}_t^\mu := \sigma(\mathscr{F}_t^{0*}\cup\mathscr{N}^\mu)$. 然后令

$$\mathscr{F}_t := \bigcap_\mu \mathscr{F}_t^\mu.$$

可以证明对关于 $\mathscr{E}^*$ 的自然流 $\mathscr{F}^{0*},\mathscr{F}_t^{0*}$ 的强化与对关于 $\mathscr{E}$ 的自然流 $\mathscr{F}^0,\mathscr{F}_t^0$ 的强化是一致的, 所以下面证明中有时经常使用后者. 另外, 后面我们说几乎处处或几乎肯定地或几乎所有样本而不具体指定概率, 总是指在 $\mathscr{N}$ 的一个事件外. 再重复一遍, 自然流与具体的过程无关, 由样本空间唯一决定.

自然的问题是如果 $B\in\mathscr{F}$, $\mathbb{P}^x(B)$ 关于 x 还可测吗? 我们有下列性质:

(1) 若 $B\in\mathscr{F}$, 则 $\mathbb{P}^\cdot(B)$ 是普遍可测的;

(2) X_t 是 $(\Omega,\mathscr{F}_t)$ 到 $(E,\mathscr{E}^*)$ 可测的;

(3) θ_t 是 $(\Omega,\mathscr{F})$ 上可测的.

这类性质的证明是标准的. 如证 (1). 设 $H\in b\mathscr{F}$, 则对任何 μ, 存在 $H_1,H_2\in\mathscr{F}^0$, 使得 $H_1\leqslant H\leqslant H_2$ 且 $\mathbb{E}^\mu(H_1)=\mathbb{E}^\mu(H_2)$. 置 $\phi_i(x):=\mathbb{E}^x(H_i)$, $i=1,2$, 那么 $\phi_1\leqslant\mathbb{E}^\cdot(H)\leqslant\phi_2$, $\phi_1,\phi_2\in b\mathscr{E}$ 且 $\mu(\phi_1)=\mu(\phi_2)$. 因此 $\mathbb{E}^\cdot(H)\in b\mathscr{E}^*$. 对于 (2), 任取 $f\in b\mathscr{E}^*$, 因为 μP_t 是概率, 故存在 $f_1,f_2\in b\mathscr{E}$, 满足 $f_1\leqslant f\leqslant f_2$ 且 $\mu P_t(f_1)=\mu P_t(f_2)$. 推出 $\mathbb{E}^\mu f_1(X_t)=\mathbb{E}^\mu f_2(X_t)$. 因为 $f_i(X_t)\in b\mathscr{F}_t^0$, 故 $f(X_t)\in b\mathscr{F}_t$. 利用单调类定理, (3) 是显然的.

定理 4.1.1 设 $Y\in b\mathscr{F}$, 则对任何 $x\in E$, $t\geqslant 0$,

$$\mathbb{E}^x(Y\circ\theta_t|\mathscr{F}_t)=\mathbb{E}^{X_t}(Y).$$

证明　由上面性质 (1) 知 $\omega \mapsto \mathbb{E}^{X_t(\omega)}(Y)$ 是 $\mathscr{F}_t$ 可测的, 因此仅需验证对任何 $\Lambda \in \mathscr{F}_t$,

$$\mathbb{E}^x(Y \circ \theta_t; \Lambda) = \mathbb{E}^x[\mathbb{E}^{X_t}(Y); \Lambda].$$

由 $\mathscr{F}_t$ 的定义, 我们只需对 $\Lambda \in \mathscr{F}_t^0$ 验证就够了. 设 μ 是 $\mathbb{P}^x$ 在 X_t 下的像, 由 $\mathscr{F}$ 的定义, 存在有界的 $\mathscr{F}^0$ 可测函数 Y_0, 使得 $\{Y \neq Y_0\} \subset \Gamma \in \mathscr{F}^0$ 且 $\mathbb{P}^\mu(\Gamma) = 0$. 因此 $\{Y \circ \theta_t \neq Y_0 \circ \theta_t\} \subset \theta_t^{-1}(\Gamma)$, 这样由 Markov 性, $\mathbb{P}^x(\theta_t^{-1}\Gamma) = \mathbb{E}^x(1_\Gamma \circ \theta_t) = \mathbb{E}^x[\mathbb{P}^{X_t}(\Gamma)] = \mathbb{P}^\mu(\Gamma) = 0$, 因此上式左边 Y 可用 Y_0 代替. 另一方面,

$$\mathbb{E}^x(\mathbb{E}^{X_t}|Y - Y_0|) = \mathbb{E}^\mu|Y - Y_0| = 0,$$

故相对于 $\mathbb{P}^x$, $\mathbb{E}^{X_t}Y = \mathbb{E}^{X_t}Y_0$ a.s., 故上式右边 Y 也可用 Y_0 代替, 而这时上式恰是定义中所述的 Markov 性. □

这个定理说明 X 关于强化了的自然流也有 Markov 性. 下面的定理给出强化了的自然流是右连续的一个充分条件.

定理 4.1.2　如果 X 关于 $(\mathscr{F}_{t+}^0)$ 也有 Markov 性, 则 $(\mathscr{F}_t)$ 满足通常条件.

证明　只需验证 $(\mathscr{F}_t)$ 是右连续的. 条件说明对任何 $Y \in b\mathscr{F}^0$ 及 $(E, \mathscr{E})$ 上有限测度 μ, 有

$$\mathbb{E}^\mu(Y \circ \theta_t | \mathscr{F}_{t+}^0) = \mathbb{E}^{X_t}(Y) = \mathbb{E}^\mu(Y \circ \theta_t | \mathscr{F}_t^0).$$

故对 $Y = f_1(X_{t_1}) \cdots f_n(X_{t_n})$, $f_1, \cdots, f_n \in b\mathscr{E}$,

$$\begin{aligned}
\mathbb{E}^\mu(Y | \mathscr{F}_{t+}^0) &= \prod_{t_i \leqslant t} f_i(X_{t_i}) \cdot E^\mu\left(\prod_{t_j > t} f_j(X_{t_j - t}) \circ \theta_t | \mathscr{F}_{t+}^0\right) \\
&= \prod_{t_i \leqslant t} f_i(X_{t_i}) \cdot \mathbb{E}^\mu\left(\prod_{t_j > t} f_j(X_{t_j - t}) \circ \theta_t | \mathscr{F}_t^0\right) \\
&= \mathbb{E}^\mu(Y | \mathscr{F}_t^0),
\end{aligned}$$

推出对任何 $Y \in b\mathscr{F}^0$, $\mathbb{E}^\mu(Y | \mathscr{F}_{t+}^0) = \mathbb{E}^\mu(Y | \mathscr{F}_t^0)$. 现在取 $A \in \mathscr{F}_{t+}^0$, 则 $1_A = \mathbb{E}^\mu(1_A | \mathscr{F}_t^0)$ a.s. $\mathbb{P}^\mu$, 即与一个 $\mathscr{F}_t^0$ 可测函数只相差一个 $\mathbb{P}^\mu$ 零测集. 因 $\mathscr{F}_t$ 是 $\mathscr{F}_t^0$ 的强化, 故 $A \in \mathscr{F}_t$, 即 $\mathscr{F}_{t+}^0 \subset \mathscr{F}_t$, 因此 $\mathscr{F}_{t+} \subset \mathscr{F}_t$. 最后的结论的理由是强化与取

右极限是可交换的: 强化后的右极限 $\mathscr{F}_{t+}$ 等于右极限 $\mathscr{F}_{t+}^0$ 的强化. □

说 T 是一个停时而不具体指明相对应的流时, 是指它是关于强化后的自然流 $(\mathscr{F}_t)$ 的停时. 现在我们来定义关于随机时间的推移算子, 设 $H:\Omega\longrightarrow\mathbf{R}_+$, 对 $\omega\in\Omega$, 且 $H(\omega)<\infty$, 定义 $\theta_H(\omega):=\theta_{H(\omega)}(\omega)$, 显然 θ_H 也是 Ω 上的映射. 对任何 $t\geqslant 0$,

$$X_t\circ\theta_H = X_{t+H}.$$

如果 H 可测, 那么 θ_H 是 $(\Omega,\mathscr{F})$ 上可测映射.

定义 4.1.2 称 X 关于一个适应的流 $(\mathscr{M}_t)$ 有强 Markov 性, 如果对任何 $(\mathscr{M}_t)$ 停时 T 及 $f\in b\mathscr{E}$, $x\in E$, 有

$$\mathbb{E}^x(f\circ X_{t+T}1_{\{T<\infty\}}|\mathscr{M}_T) = 1_{\{T<\infty\}}\mathbb{E}^{X(T)}(f\circ X_t). \tag{4.1.1}$$

如果 X 关于右极限流 $(\mathscr{F}_{t+})$ 有强 Markov 性, 我们说 X 有强 Markov 性或者 X 是强 Markov 过程.

单调类方法可以证明 X 关于一个适应的流 $(\mathscr{M}_t)$ 有强 Markov 性当且仅当对任何 $Y\in b\mathscr{F}^0$, $x\in E$ 及 $(\mathscr{M}_t)$ 停时 T, 有

$$\mathbb{E}^x(Y\circ\theta_T 1_{\{T<\infty\}}|\mathscr{M}_T) = 1_{\{T<\infty\}}\mathbb{E}^{X_T}(Y).$$

注意从强 Markov 性容易看出 X 关于右极限流 $(\mathscr{F}_{t+})$ 是 Markov 的. 因此强 Markov 过程的强化后的自然流满足通常条件. 下面证明定义中的强 Markov 性等价于下面看上去更为简单的形式.

定理 4.1.3 过程 X 关于一个适应的流 $(\mathscr{M}_t)$ 有强 Markov 性当且仅当对任何 $(\mathscr{M}_t)$ 停时 T, $f\in C_u(E)$ 及 $x\in E$, $t\geqslant 0$, 下式成立

$$\mathbb{E}^x(f\circ X_{t+T};T<\infty) = \mathbb{E}^x(\mathbb{E}^{X(T)}f(X_t);T<\infty). \tag{4.1.2}$$

证明 只需验证充分性. 由单调类定理, 只需验证 (4.1.1) 对于 $f\in C_u(E)$ 成立就足够了. 任取 $\Lambda\in\mathscr{M}_T$, 定义 $T_\Lambda := T\cdot 1_\Lambda + \infty\cdot 1_{\Lambda^c}$, 那么容易验证 T_Λ 也是 $(\mathscr{M}_t)$ 停时, 代入上式得

$$\mathbb{E}^x(f(X_{t+T_\Lambda});T_\Lambda<\infty) = \mathbb{E}^x(f(X_{t+T});\Lambda,T<\infty),$$

$$\mathbb{E}^x(\mathbb{E}^{X(T_\Lambda)}f(X_t);T_\Lambda<\infty) = \mathbb{E}^x(\mathbb{E}^{X(T)}f(X_t);\Lambda,T<\infty),$$

由定义 4.1.2 推出 X 关于 $(\mathscr{M}_t)$ 具有强 Markov 性. □

引理 4.1.2 (Dynkin) 设 X 是强 Markov 的, T 是停时, $\alpha > 0$, $f \in b\mathscr{E}^*$, 则

$$\mathbb{E}^x \mathrm{e}^{-\alpha T} U^\alpha f(X_T) = \mathbb{E}^x \int_T^\infty \mathrm{e}^{-\alpha t} f(X_t) \mathrm{d}t.$$

证明 由强 Markov 性,

$$\begin{aligned}
&\mathbb{E}^x \mathrm{e}^{-\alpha T} U^\alpha f(X_T) \\
&= \mathbb{E}^x \left(\mathrm{e}^{-\alpha T} \mathbb{E}^{X_T} \int_0^\infty \mathrm{e}^{-\alpha t} f(X_t) \mathrm{d}t \right) \\
&= \mathbb{E}^x \left(\mathrm{e}^{-\alpha T} \left(\int_0^\infty \mathrm{e}^{-\alpha t} f(X_t) \mathrm{d}t \right) \circ \theta_T \right) \\
&= \mathbb{E}^x \int_0^\infty \mathrm{e}^{-\alpha (t+T)} f(X_{t+T}) \mathrm{d}t \\
&= \mathbb{E}^x \int_T^\infty \mathrm{e}^{-\alpha t} f(X_t) \mathrm{d}t.
\end{aligned}$$

完成证明. □

借助豫解算子, 我们来叙述右假设 2.

定义 4.1.3 设转移半群 (P_t) 和过程 X 满足(HD1). 说 (P_t) 满足右假设 2(HD2), 如果对任何 $f \in C_u(E)$, $\alpha > 0$, 过程 $t \mapsto U^\alpha f(X_t)$ 的几乎所有轨道是右连续的.

如果满足 (HD1) 的半群 (P_t) 的豫解 U^α 把连续函数映为连续函数, 则 (HD2) 成立. 一个满足两个右假设 (HD1), (HD2) 的转移半群 (P_t) 和对应的随机过程称为是右半群和右 Markov 过程, 简称右过程, 我们通常说随机过程

$$X = (\Omega, \mathscr{F}, \mathscr{F}_t, X_t, \theta_t, \mathbb{P}^x)$$

是转移半群为 (P_t), 状态空间为 E 的右过程. 一个右过程称为是 Borel 的, 如果 E 是 Lusin 空间(是紧度量空间的 Borel 子集)且 (P_t) 是 Borel 转移半群. 上面我们看到 (HD1) 实际上是右连续假设, 而下面我们将证明 (HD2) 差不多就是强 Markov 性假设.

定理 4.1.4 假设右假设 1 成立, 那么 (HD2) 蕴含着过程 (X_t) 是强 Markov 过程.

证明　由 (HD2), 对 $f \in C_u(E)$, 过程 $U^\alpha f(X.)$ 是右连续的. 设 T 是 $(\mathscr{F}_{t+})$ 停时, 不妨设 $T < \infty$. 令 T_n 是 T 的离散化,

$$T_n = \sum_{k=1}^{\infty} \frac{k}{2^n} 1_{\{\frac{k-1}{2^n} \leqslant T < \frac{k}{2^n}\}},$$

则

$$\left\{T_n = \frac{k}{2^n}\right\} = \left\{T < \frac{k}{2^n}\right\} \Big\backslash \left\{T < \frac{k-1}{2^n}\right\} \in \mathscr{F}_{\frac{k}{2^n}}.$$

因此由轨道的右连续性及 X 关于流 $(\mathscr{F}_t)$ 的简单 Markov 性,

$$\begin{aligned}
&\mathbb{E}^x \int_0^\infty \mathrm{e}^{-\alpha t} f(X_{t+T}) \mathrm{d}t = \lim_n \mathbb{E}^x \int_0^\infty \mathrm{e}^{-\alpha t} f(X_{t+T_n}) \mathrm{d}t \\
&= \lim_n \sum_k \mathbb{E}^x \left(\int_0^\infty \mathrm{e}^{-\alpha t} f(X_{t+\frac{k}{2^n}}) \mathrm{d}t; T_n = \frac{k}{2^n} \right) \\
&= \lim_n \sum_k \mathbb{E}^x \left(\mathbb{E}^{X(\frac{k}{2^n})} \int_0^\infty \mathrm{e}^{-\alpha t} f(X_t) \mathrm{d}t; T_n = \frac{k}{2^n} \right) \\
&= \lim_n \mathbb{E}^x U^\alpha f(X_{T_n}) = \mathbb{E}^x U^\alpha f(X_T) \\
&= \int_0^\infty \mathrm{e}^{-\alpha t} \mathbb{E}^x \mathbb{E}^{X(T)} f(X_t) \mathrm{d}t,
\end{aligned}$$

即 $t \mapsto \mathbb{E}^x f(X_{t+T})$ 与 $t \mapsto \mathbb{E}^x \mathbb{E}^{X(T)} f(X_t)$ 有相同的 Laplace 变换, 因为两者都是右连续的, 故对任何 t,

$$\mathbb{E}^x f(X_{t+T}) = \mathbb{E}^x \mathbb{E}^{X_T} f(X_t).$$

推出 X 是强 Markov 的. □

实际上, 反过来也差不多成立. 如果 (P_t) 是 Borel 转移半群, 那么 X 是强 Markov 过程蕴含着 (HD2) 成立. 一般地, 如果没有 Borel 转移半群的假设, 一个具有强 Markov 性的右连续简单 Markov 过程不一定是右过程.

定理 4.1.5　如果 X 是简单右连续 Markov 过程, 具有强 Markov 性且 (P_t) 是 Borel 转移半群, 那么 (HD2) 成立.

证明 只需证明如果 $f \in b\mathscr{E}$, 那么 $t \mapsto U^q f(X_t)$ 是右连续的. 因为 P_t 是 Borel 的, 故 $U^q f \in b\mathscr{E}$. 因此 $t \mapsto U^q f(X_t)$ 是可选过程. 由 Dynkin 公式对任何停时列 $T_n \downarrow T$,

$$\mathbb{E}^x(\mathrm{e}^{-qT_n} U^q f(X_{T_n})) = \mathbb{E}^x \int_{T_n}^{\infty} \mathrm{e}^{-qt} f(X_t)\mathrm{d}t$$

$$\longrightarrow \mathbb{E}^x \int_{T}^{\infty} \mathrm{e}^{-qt} f(X_t)\mathrm{d}t = \mathbb{E}^x(\mathrm{e}^{-qT} U^q f(X_T)),$$

由 Meyer 的截面定理推出过程 $t \mapsto \mathrm{e}^{-qt}U^q f(X_t)$ 是右连续的, 因此过程 $t \mapsto U^q f(X_t)$ 也是右连续的. □

下面的定理通常称为 Blumenthal 0-1 律.

定理 4.1.6 (Blumenthal) 设 $A \in \mathscr{F}_0$, 则对任何 $x \in E$, $\mathbb{P}^x(A)$ 等于 0 或 1.

证明 容易验证对任何 $B \in \mathscr{F}^0$, $\theta_0^{-1}B = B$. 现在 $A \in \mathscr{F}_0$, 则对任何 E 上的分布 μ, 存在 $B \in \mathscr{F}_0^0$ 使得 $\mathbb{P}^\mu(B \triangle A) = 0$. 由 Markov 性, $\mathbb{P}^\mu(\theta_0^{-1}B \triangle \theta_0^{-1}A) = \mathbb{P}^\mu(B \triangle A) = 0$, 因此 $\mathbb{P}^\mu(A \triangle \theta_0^{-1}A) = \mathbb{P}^\mu(B \triangle \theta_0^{-1}B) = 0$ 对任何 μ 成立. 再利用 Markov 性,

$$\mathbb{P}^x(A) = \mathbb{P}^x(A \cap \theta_0^{-1}A) = \mathbb{E}^x(1_A \circ \theta_0; A) = \mathbb{E}^x[\mathbb{P}^{X_0}(A); A] = (\mathbb{P}^x(A))^2,$$

因此定理结论成立. □

例 4.1.1 我们给一个连续简单 Markov 过程但不满足强 Markov 性的例子. 设 $E = [0, \infty)$, 概率空间 $(\Omega', \mathscr{F}', \mathbb{P}')$ 上随机变量 T' 服从参数为 $c > 0$ 的指数分布. 定义对任何 $x > 0$, $X_t^x := x + t$, 而

$$X_t^0 = \begin{cases} 0, & 0 \leqslant t < T', \\ t - T', & t \geqslant T'. \end{cases}$$

取 Ω 为连续轨道样本空间, 定义 $\xi_x \omega'(t) := X_t^x(\omega')$, 那么 ξ_x 是 Ω' 到 Ω 的可测映射. 定义 $\mathbb{P}^x := \mathbb{P}' \circ \xi_x^{-1}$, 则 $\{\mathbb{P}^x\}$ 是 Markov 的但不是强 Markov 的.

当 $x > 0$ 时, $X_t = x + t$ a.s. $\mathbb{P}^x$. 设 T 是 $\{0\}$ 的首离时(0 的 holding time), 那么 T 服从参数 c 的指数分布, 且在概率 $\mathbb{P}^0$ 下,

$$X_t = \begin{cases} 0, & 0 \leqslant t < T, \\ t - T, & t \geqslant T. \end{cases}$$

令 $P_t(x, A) := \mathbb{P}^x(X_t \in A)$, 那么

$$P_t(x, \mathrm{d}y) = \begin{cases} \epsilon_{x+t}(\mathrm{d}y), & x > 0, \\ \mathbb{P}^0(T > t)\epsilon_0(\mathrm{d}y) + \mathbb{P}^0(t - T \in \mathrm{d}y, T < t), & x = 0. \end{cases}$$

让我们验证对于 $f, g \in b\mathscr{E}$,

$$\mathbb{E}^x(f(X_{t+s})g(X_s)) = P_s(gP_t f)(x).$$

事实上, 当 $x > 0$ 时显然. 设 $x = 0$, 写 $\mathbb{P} = \mathbb{P}^0$, 则

$$\begin{aligned}
&\mathbb{E}^0(g(X_s)f(X_{t+s})) \\
&= \mathbb{E}^0(g(X_s)f(X_{t+s}), T < s) + \mathbb{E}^0(g(X_s)f(X_{t+s}), s < T < t + s) \\
&\quad + \mathbb{E}^0(g(X_s)f(X_{t+s}), T > t + s) \\
&= \mathbb{E}^0(g(s - T)f(t + s - T), T < s) + g(0)\mathbb{E}^0(f(t + s - T), s < T < t + s) \\
&\quad + g(0)f(0)\mathbb{P}^0(T > t + s) \\
&= \mathbb{E}^0(g(s - T)f(t + s - T), T < s) + g(0)\mathbb{P}^0(T > s)\mathbb{E}^0(f(t - T), T < t) \\
&\quad + g(0)f(0)\mathbb{P}^0(T > t + s) \\
&= \mathbb{P}^0(g(s - T)f(s + t - T), T < s) \\
&\quad + \mathbb{P}(T > s)g(0)\left(\mathbb{P}^0(f(t - T), T < t) + \mathbb{P}^0(T > t)f(0)\right) \\
&= \mathbb{P}^0(g(s - T)P_t f(s - T), T < s) + \mathbb{P}^0(T > s)g(0)P_t f(0) = P_s(gP_t f)(0).
\end{aligned}$$

第三个等号因为 T 是指数分布,

$$\mathbb{E}^0(f(t + s - T), s < T < t + s) = \int_s^{t+s} c\mathrm{e}^{-cu} f(t + s - u)\mathrm{d}u$$

$$= \mathrm{e}^{-cs}\int_0^t c\mathrm{e}^{-cu} f(t - u)\mathrm{d}u = \mathbb{P}^0(T > s)\mathbb{E}^0(f(t - T), T < t).$$

第四个等号也因为 T 是指数分布的遗忘性 $\mathbb{P}^0(T > t+s) = \mathbb{P}^0(T > t)\mathbb{P}^0(T > s)$. 因此 $\{\mathbb{P}^x\}$ 是 Markov 的, 它显然不是强 Markov 的. 对 $t > 0$, $\mathbb{P}^0(X_{t+T} = 0) = 0$, 而 $\mathbb{P}^0(\mathbb{P}^{X_T}(X_t = 0)) = \mathbb{P}^0(X_t = 0) = \mathrm{e}^{-ct}$.

在第二章中引入的 Poisson 过程和 Brown 运动都是 Borel 右过程. 设 B 是 d-维 Brown 运动, $A \in \mathscr{B}(\mathbf{R}^d)$, τ 是 A 的首中时, 那么

$$P_t^A f(x) := \mathbb{E}^x(f(B_t); \tau > t),\ x \in \mathbf{R}^d,\ t \geqslant 0$$

定义了另一个 $\mathbf{R}^d$ 上的转移半群, 但是它不是 Borel 转移半群, 因为 $\{\tau > t\} \in \mathscr{F}_t$ 一般不在 $\mathscr{F}_t^0$ 中, 故 $P_t^A f$ 仅是 $\mathscr{E}^*$ 可测的, 即使 f 是 $\mathscr{E}$-可测的. 因此右半群或者右过程的引入是必要的.

回到豫解算子. 如果右半群 (P_t^1) 与 (P_t^2) 有相同的豫解, 由 Laplace 变换的唯一性推出对任何有界连续函数 f, $P_t^1 f = P_t^2 f$, 因为两者关于 t 都是右连续的. 因此 $P_t^1 = P_t^2$, 即豫解唯一地决定半群本身, 但位势算子最主要的性质是下面的豫解方程.

定理 4.1.7 对任何 $\alpha, \beta > 0$, 有

$$U^\alpha - U^\beta + (\alpha - \beta) U^\alpha U^\beta = 0.$$

故位势算子可以交换 $U^\alpha U^\beta = U^\beta U^\alpha$.

证明　由半群性质 $P_t P_s = P_{t+s}$ 容易验证

$$\mathrm{e}^{-\alpha s} P_s U^\alpha f = \int_s^\infty \mathrm{e}^{-\alpha t} P_t f \mathrm{d}t.$$

因此

$$\begin{aligned} U^\beta U^\alpha f(x) &= \int_0^\infty \mathrm{e}^{-\beta s} P_s U^\alpha f(x) \mathrm{d}s \\ &= \int_0^\infty \mathrm{e}^{(\alpha-\beta)t} \mathrm{d}s \int_s^\infty \mathrm{e}^{-\alpha t} P_t f(x) \mathrm{d}t \\ &= \int_0^\infty \mathrm{e}^{-\alpha t} P_t f(x) \mathrm{d}t \int_0^t \mathrm{e}^{(\alpha-\beta)s} \mathrm{d}s \\ &= \frac{U^\beta f(x) - U^\alpha f(x)}{\alpha - \beta}. \end{aligned}$$

□

下面我们介绍生成算子. 不妨设 X 是 Borel 右过程. 用 $\mathscr{E}_b$ 表示 E 上有界可测函数全体, $\|f\| := \sup_{x\in E}|f(x)|$, $f \in \mathscr{E}_b$, 那么 $\|P_tf\| \leqslant \|f\|$, 即 P_t 是 Banach 空间 $\mathscr{E}_b$ 上压缩算子. 让我们考虑形式算子

$$A := \lim_{t\to 0}\frac{P_t - I}{t} = \frac{\mathrm{d}P_t}{\mathrm{d}t}|_{t=0},$$

则

$$\frac{\mathrm{d}P_t}{\mathrm{d}t} = \lim_{s\to 0}\frac{P_{t+s} - P_t}{s} = AP_t.$$

因此 $P_t = \mathrm{e}^{tA}$ 且

$$U^\alpha = \int_0^\infty \mathrm{e}^{-\alpha t}P_t\mathrm{d}t = \int_0^\infty \mathrm{e}^{-(\alpha-A)}\mathrm{d}t = (\alpha - A)^{-1},$$

或者说 $A = \alpha - (U^\alpha)^{-1}$. 这个算子 A 称为是 (P_t) 的生成算子或无穷小算子. 当半群满足一定条件时, 这个形式算子有一个精确的定义并唯一地决定半群, 称为是 Hille-Yosida 理论. 后面介绍的 Feller 半群在适当选取的空间后是满足这些条件的. 这里半群一般不满足这些条件, 但我们仍然可以给 A 一个严格的定义.

定义 $\mathscr{R}_\alpha := U^\alpha\mathscr{E}_b$, $\mathscr{N}_\alpha := (U^\alpha)^{-1}(\{0\})$, 分别是值域与零集.

引理 4.1.3 $\mathscr{R}_\alpha$ 与 $\mathscr{N}_\alpha$ 与 α 无关且 $\mathscr{R}_\alpha \cap \mathscr{N}_\alpha = \{0\}$.

证明 对任何 $f \in \mathscr{E}_b$, 由豫解方程 $U^\alpha f = U^\beta[f - (\alpha - \beta)U^\alpha f]$, 因此 $\mathscr{R}_\alpha \subset \mathscr{R}_\beta$. 同理 $\mathscr{R}_\beta \subset \mathscr{R}_\alpha$, 推出 $\mathscr{R}_\alpha = \mathscr{R}_\beta$. 类似地, 由于 $U^\alpha f = (I - (\alpha - \beta)U^\alpha)U^\beta f$, 故 $U^\beta f = 0$ 蕴含 $U^\alpha f = 0$, 因此 $\mathscr{N}_\beta \subset \mathscr{N}_\alpha$. 同理推出相反的包含关系.

写 $\mathscr{R} = \mathscr{R}_\alpha$, $\mathscr{N} = \mathscr{N}_\alpha$. 如果 $f \in \mathscr{R} \cap \mathscr{N}$, 那么存在 $g \in \mathscr{E}_b$, 使得 $f = U^1 g$ 且对任何 β, $U^\beta f = 0$. 因为 $\mathrm{e}^{-s}P_sf(x) = \int_s^\infty \mathrm{e}^{-t}P_tg\mathrm{d}t$, 故 $\lim_{s\to 0}P_sf(x) = f(x)$. 而 $U^\beta f(x) = 0$ 蕴含 $s \mapsto P_sf(x)$ 几乎处处为零, 因此 $f(x) = 0$. □

现在定义 $D(A) := \mathscr{R}$, 对任何 $f \in D(A)$, 定义

$$Af := \alpha f - g,$$

其中 $U^\alpha g = f$. 当然 g 不是唯一的, 但两个这样的 g 的差在 $\mathscr{N}$ 中, 因此 Af 在商空间 $\mathscr{E}_b/\mathscr{N}$ 上唯一决定的. 现在需要证明右边与 α 无关, 即验证 $\alpha f - g - (\beta f - h) \in \mathscr{N}$,

其中 $U^\beta h = f$. 由豫解方程

$$\begin{aligned}
&U^\alpha(\alpha f - g - \beta f + h) \\
&= \alpha U^\alpha f - f - \beta U^\alpha f + U^\alpha h \\
&= (\alpha - \beta)U^\alpha f - f + (U^\beta - (\alpha - \beta)U^\alpha U^\beta)h \\
&= 0.
\end{aligned}$$

这里的 $(A, D(A))$ 称为是 X 的生成算子.

定理 4.1.8 设 $X^i = (\Omega, \mathscr{F}^0, X_t, \mathbb{P}_i^x)$, $i = 1, 2$ 是两个 Markov 过程, 生成算子分别为 $A_i : i = 1, 2$. 如果 $A_1 = A_2$, 那么 $\mathbb{P}_1^x = \mathbb{P}_2^x$, $x \in E$.

证明 $A_1 = A_2$ 等价于两个豫解 U_1^α 与 U_2^α 的值域 $\mathscr{R}$ 与零集 $\mathscr{N}$ 分别相同. 现在需要验证对任何 $g \in \mathscr{E}_b$, $U_1^\alpha g = U_2^\alpha g$. 因为值域相同, 存在 $f \in \mathscr{E}_b$, 使得 $U_1^\alpha g = U_2^\alpha f$, 记为 u, 那么 $A_1 u = \alpha u - g$, $A_2 u = \alpha u - f$, 推出 $f - g \in \mathscr{N}$. 因此 $U_1^\alpha g = U_2^\alpha f = U_2^\alpha g$. □

例 4.1.2 设 $\{\mathbb{P}^x\}$ 是 $\mathbf{R}$ 上 Brown 运动. 这时 Ω 是连续轨道全体. 我们来算它的生成算子 A. 任取有界可测函数 f, $\alpha > 0$. 令 $u = U^\alpha f$, 那么由 Brown 运动的位势算子的表达式(见习题)

$$\begin{aligned}
u(x) &= \int_{\mathbf{R}} f(y + x)\mathrm{d}y \int_0^\infty \mathrm{e}^{-\alpha t} \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{y^2}{2t}} \mathrm{d}t \\
&= \int_{\mathbf{R}} f(y + x) \frac{1}{\sqrt{2\alpha}} \mathrm{e}^{-\sqrt{2\alpha}|y|} \mathrm{d}y \\
&= \int_{-\infty}^\infty \frac{1}{\sqrt{2\alpha}} \mathrm{e}^{-\sqrt{2\alpha}|y - x|} f(y)\mathrm{d}y \\
&= \frac{\mathrm{e}^{-\sqrt{2\alpha}x}}{\sqrt{2\alpha}} \int_{-\infty}^x \mathrm{e}^{\sqrt{2\alpha}y} f(y)\mathrm{d}y + \frac{\mathrm{e}^{\sqrt{2\alpha}x}}{\sqrt{2\alpha}} \int_x^\infty \mathrm{e}^{-\sqrt{2\alpha}y} f(y)\mathrm{d}y.
\end{aligned}$$

因此 u 是绝对连续的, 且

$$u'(x) = \mathrm{e}^{\sqrt{2\alpha}x} \int_x^\infty \mathrm{e}^{-\sqrt{2\alpha}y} f(y)\mathrm{d}y - \mathrm{e}^{-\sqrt{2\alpha}x} \int_{-\infty}^x \mathrm{e}^{\sqrt{2\alpha}y} f(y)\mathrm{d}y.$$

显然 u' 也是绝对连续的, 且 $u'' = 2\alpha u - 2f$. 令

$$\mathscr{R}_+ := \{u \in b\mathscr{B}(\mathbf{R}):\ u\text{绝对连续},\ u'\ \text{绝对连续且}\ u''\ \text{有界}\},$$

那么 $\mathscr{R} \subset \mathscr{R}_+$. 反之, 若 $u \in \mathscr{R}_+$, 令 $f = \alpha u - \frac{1}{2}u''$, 那么 $f \in b\mathscr{B}(\mathbf{R})$ 且 $v := U^\alpha f$ 也满足 $v'' = 2\alpha v - 2f$. 因此 $w := u - v$ 满足方程 $w'' = 2w$ 且有界, 因此 $u = v$. 故而 $\mathscr{R}_+ \subset \mathscr{R}$, 推出 $\mathscr{R} = \mathscr{R}^+$. 然后我们再来看 $\mathscr{N}$. 如果 $u = U^\alpha f = 0$, 那么由上面方程得 $f = 0$ a.e., 因此

$$\mathscr{N} = \{f \in b\mathscr{B}(\mathbf{R}):\ f = 0\ a.e.\}.$$

这样我们得到 $D(A) = \mathscr{R}$ 且 $Au = \frac{1}{2}u''$. ▮

例 4.1.3 (反射 Brown 运动) 设 $\{\mathbb{P}^x\}$ 是 $\mathbf{R}$ 上 Brown 运动. 考虑一种状态空间的变换. 用 $\hat{\Omega}$ 表示以 $[0,\infty)$ 为状态空间的样本空间, $\hat{\mathscr{F}}^0, \hat{\mathscr{F}}^0_t$ 是其对应 σ- 代数, 定义

$$\xi\omega(t) := |\omega(t)|,\ \omega \in \Omega,\ t \geqslant 0.$$

它是 Ω 到 $\hat{\Omega}$ 的可测映射. 对任何 $x \geqslant 0$, 定义

$$\hat{\mathbb{P}}^x := \mathbb{P}^x \circ \xi^{-1}.$$

下面我们证明 $\{\hat{\mathbb{P}}^x : x \geqslant 0\}$ 也是 Markov 过程. 先证明 $\mathbb{P}^{-x}\circ\xi^{-1} = \mathbb{P}^x\circ\xi^{-1}$. 事实上, $X_t\circ\xi = |X_t|$ 且因为热核半群的对称性 $P_t(x,A) = P_t(-x,-A)$, 故

$$\mathbb{P}^{-x}(|X_t| \in A) = \mathbb{P}^{-x}(X_t \in A \cup (-A)) = \mathbb{P}^x(X_t \in A \cup (-A)) = \mathbb{P}^x(|X_t| \in A).$$

由单调类定理推出 $\mathbb{P}^x$ 与 $\mathbb{P}^{-x}$ 在 ξ 有相同的像.对任何 $t,s \geqslant 0$, $A \subset \mathscr{B}(\mathbf{R}_+)$, $\Lambda \in \hat{\mathscr{F}}^0_t$, $x \geqslant 0$, 那么 $\xi^{-1}\Lambda \in \mathscr{F}_t$, 且

$$\begin{aligned}
\hat{\mathbb{P}}^x(X_{t+s} \in A, \Lambda) &= \mathbb{P}^x(|X_{t+s}| \in A, \xi^{-1}\Lambda) \\
&= \mathbb{E}^x(\mathbb{P}^{X_t}(|X_s| \in A), \xi^{-1}\Lambda) \\
&= \mathbb{E}^x(\hat{\mathbb{P}}^{|X_t|}(X_s \in A), \xi^{-1}\Lambda) \\
&= \hat{\mathbb{E}}^x(\hat{\mathbb{P}}^{X_t}(X_s \in A), \Lambda).
\end{aligned}$$

过程 $\{\hat{\mathbb{P}}^x\}$ 称为是 0 点反射的 Brown 运动. 这里用到 Brown 运动的对称性.

零点反射的 Brown 运动的生成算子可利用 Brown 运动的生成算子计算. 用 ˆ· 表示对应的半群, 豫解, 生成算子. 对 $[0,\infty)$ 上的函数 f, $\hat{f}(x) := f(|x|)$, 那么 $\hat{P}_t f = P_t\hat{f}$, $\hat{U}^\alpha f = U^\alpha \hat{f}$. 因此 $D(\hat{A}) = \{U^\alpha \hat{f} : f \in b\mathscr{B}(\mathbf{R}_+)\}$. 由上例得知

$$D(\hat{A}) = \{u \in b\mathscr{B}(\mathbf{R}_+) : \ u \text{ 绝对连续}, \ u' \text{ 绝对连续且 } u'' \text{ 有界}, \ u'_+(0) = 0\},$$

且 $\hat{A}u = \dfrac{1}{2}u''$. ∎

例 4.1.4 上例由 Brown 运动得到反射 Brown 运动的程序是一种称为状态空间变换的特例. 设 $\{\mathbb{P}^x\}$ 是状态空间 E 上的 Markov 过程, $\hat{E}$ 是另外一个 Lusin 空间, ψ 是 E 到 $\hat{E}$ 的连续满射, 那么 ψ 诱导 E 的样本空间 Ω 到 $\hat{E}$ 的样本空间的映射, 仍记为 ψ: $\psi\omega(t) = \psi(\omega(t))$. 但它要诱导 $\hat{E}$ 上的 Markov 过程还需要一个苛刻的条件: 对任何 $x, y \in E$ 及 $\hat{\Omega}$ 的有界随机变量 $\hat{H}$, 如果 $\psi(x) = \psi(y)$, 那么

$$\mathbb{E}^x(\hat{H}\circ\psi) = \mathbb{E}^y(\hat{H}\circ\psi).$$

这时我们定义 $\hat{\mathbb{P}}^{\psi(x)} = \mathbb{P}^x\circ\psi^{-1}$, $x \in E$. 可以证明 $\{\hat{\mathbb{P}}^{\hat{x}} : \hat{x} \in \hat{E}\}$ 是 Markov 过程. 用 $(\hat{P}_t)$ 表示对应的转移半群, 则 $(\hat{P}_t\hat{f})\circ\psi = P_t(\hat{f}\circ\psi)$, 其中 $\hat{f}$ 是 $\hat{E}$ 上的非负可测函数. ∎

附 σ- 代数的完备化与强化 在这里我们简要介绍并讨论 σ- 代数的完备化与强化的概念. 设 $(E,\mathscr{E})$ 是一个可测空间. 为方便, 记号 $f \in \mathscr{E}$, $b\mathscr{E}$, $\mathscr{E}_+$ 分别表示 f 是 $(E,\mathscr{E})$ 上可测, 有界可测, 非负可测函数. 用 M 表示 $(E,\mathscr{E})$ 上有限测度全体. 让我们回忆一下完备与完备化的概念. 取 $\mu \in M$, 说 $\mathscr{E}$ 的子 σ- 代数 $\mathscr{B}$ 关于 μ 是完备的, 如果 $\mathscr{B}$ 的 μ-零测集的子集也属于 $\mathscr{B}$. 任何子 σ- 代数都可以完备化, $\mathscr{B}$ 关于 μ 的完备化, 记为 $\mathscr{B}^\mu$, 是 $\mathscr{B}$ 与 $\mathscr{B}$ 中 μ-零测集的子集所生成的 σ- 代数, 即

$$\mathscr{B}^\mu := \sigma(\mathscr{B} \cup N^\mu_{\mathscr{B}}),$$

其中 $N^\mu_{\mathscr{B}}$ 是 $\mathscr{B}$ 中的 μ-零测集及其子集全体. 显然, 如果 $\mathscr{B}_1 \subset \mathscr{B}_2$, 那么 $\mathscr{B}_1^\mu \subset \mathscr{B}_2^\mu$. 不难验证 $B \in \mathscr{B}^\mu$ 当且仅当存在 $B_1, B_2 \in \mathscr{B}$, 使得 $B_1 \subset B \subset B_2$ 且 $\mu(B_2 \setminus B_1) = 0$. 由此可以推出 $f \in b\mathscr{B}^\mu$ 当且仅当存在 $f_1, f_2 \in b\mathscr{B}$, 使得 $f_1 \leqslant f \leqslant f_2$ 且 $\mu(f_1 \neq f_2) = 0$.

例 4.1.5 设 $\mathscr{F}$ 是 $\Omega = [0,1]$ 上的 Borel σ- 代数, μ 是其上 Lebesgue 测度, 则

$(\Omega, \mathscr{F}, \mu)$ 不是完备的, 且 $\mathscr{F}$ 关于 μ 的完备化 是 Ω 上 Lebesgue 可测集全体. 平凡 σ- 代数 $\{\varnothing, \Omega\}$ 是完备的. 取 $A \in \mathscr{F}$, A 及 A^c 都至少有两个点, 则由 A 生成的 σ- 代数 $\{\varnothing, A, A^c, \Omega\}$ 当 $0 < \mu(A) < 1$ 是完备的, 否则不是完备的.

取 $x \in \Omega$, 显然 $\mathscr{F}$ 关于 Dirac 测度 ϵ_x 的完备化是 Ω 的子集全体, 而 $\{\varnothing, \Omega\}$ 关于 ϵ_x 也是完备的. ▮

用 $\mathscr{N}^\mu$ 表示 $\mathscr{E}$ 的 μ-零测集及其子集全体. $\mathscr{E}$ 的子 σ- 代数称为是增强的, 如果 $\mathscr{B} \supset \mathscr{N}^\mu$. 类似地, $\mathscr{B}$ 关于 μ 与 $\mathscr{E}$ 的强化, 记为 $\tilde{\mathscr{B}}^\mu$, 定义为 $\mathscr{B}$ 与 $\mathscr{N}^\mu$ 一起生成的 σ- 代数, 即

$$\tilde{\mathscr{B}}^\mu := \sigma(\mathscr{B} \cup \mathscr{N}^\mu).$$

显然, $\tilde{\mathscr{E}}^\mu = \mathscr{E}^\mu$, 但一般地 $\mathscr{B}^\mu \subset \tilde{\mathscr{B}}^\mu \subset \mathscr{E}^\mu$ 且如果 $\mathscr{B}_1 \subset \mathscr{B}_2$, 那么 $\tilde{\mathscr{B}}_1^\mu \subset \tilde{\mathscr{B}}_2^\mu$. 强化也有一个等价条件: $\tilde{\mathscr{B}}^\mu$ 是所有包含 $\mathscr{B}$ 且关于 μ 与 $\mathscr{E}$ 是增强的 σ- 代数的交.

例 4.1.6 仍取例 4.1.4 中的 $(\Omega, \mathscr{F}, \mu)$, 则平凡 σ- 代数 $\{\varnothing, \Omega\}$ 关于 μ 与 $\mathscr{F}$ 的强化是 Lebesgue 零测集及其补集全体, 而关于 ϵ_x 与 $\mathscr{F}$ 的强化是 Ω 子集全体. ▮

例 4.1.6 说明 $\tilde{\mathscr{B}}^\mu \neq \{B \cup N : B \in \mathscr{B}, N \in \mathscr{N}^\mu\}$, 因为这时右边一般不是 σ- 代数. 但有下列性质.

引理 4.1.4 $A \in \tilde{\mathscr{B}}^\mu$ 当且仅当存在 $B \in \mathscr{B}$, 使得 $A \triangle B \in \mathscr{N}^\mu$. 函数 f 关于 $\tilde{\mathscr{B}}^\mu$ 可测当且仅当存在 $g \in \mathscr{B}$, 使得 $\{f \neq g\} \in \mathscr{N}^\mu$.

证明　令

$$\mathscr{A} := \{A \subset \Omega : \text{存在 } B \in \mathscr{B} \text{ 使得 } A \triangle B \in \mathscr{N}^\mu\}.$$

若 $A \in \mathscr{A}$, 则 $A = [B \setminus (B \setminus A)] \cup (A \setminus B) \in \tilde{\mathscr{B}}^\mu$, 故 $\mathscr{A} \subset \tilde{\mathscr{B}}^\mu$. 反过来, $\mathscr{A}$ 显然包含了 $\mathscr{B}$ 与 $\mathscr{N}^\mu$, 而容易验证 $\mathscr{A}$ 是一个 σ- 代数, 故 $\mathscr{A} \supset \tilde{\mathscr{B}}^\mu$. 第二部分作为练习留给读者. □

习　题

1. Radon 空间的普遍可测集仍然是 Radon 空间.
2. 证明强化后的右极限 $\mathscr{F}_{t+}$ 等于右极限 $\mathscr{F}_{t+}^0$ 的强化. 用 $\tilde{\mathscr{F}}_t$ 表示 $\mathscr{F}_{t+}^0$ 的强化. 对任何 $s > t$, $\mathscr{F}_s \supset \mathscr{F}_{t+}^0$, 因此 $\mathscr{F}_{t+} \supset \tilde{\mathscr{F}}_t$. 反之, $\mathscr{F}_{t+} = \bigcap_n \mathscr{F}_{t_n}$ 其中 $t_n > t$,

$t_n \downarrow t$. 对任何 $Y \in b\mathscr{F}_{t+}$, 有 $Y \in \mathscr{F}_{t_n}$, 任取 μ, 存在 $Y_{n,1}, Y_{n,2} \in b\mathscr{F}^0_{t_n}$, 满足 $Y_{n,1} \leqslant Y \leqslant Y_{n,2}$ 且 $\mathbb{E}^\mu Y_{n,1} = \mathbb{E}^\mu Y_{n,2}$. 令 $Y_1 := \limsup Y_{n,1}$, $Y_2 := \liminf Y_{n,2}$, 那么 $Y_1, Y_2 \in b\mathscr{F}^0_{t+}$, $Y_1 \leqslant Y \leqslant Y_2$ 且 $\mathbb{E}^\mu(Y_2 - Y_1) = 0$. 因此 $Y \in \tilde{\mathscr{F}}_t$.

3. 设 T 是 $(\mathscr{F}^0_t)$ 停时, 证明: $\mathscr{F}^0_T$ 的强化是 $\mathscr{F}_T = \{A \in \mathscr{F} : A \cap \{T \leqslant t\} \in \mathscr{F}_t$ 对所有 $t\}$.
4. 求例 4.1.1 中过程 $\{\mathbb{P}^x\}$ 的生成算子.
5. 证明: 对 $\alpha > 0$,

$$\int_0^\infty \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\alpha t - \frac{x^2}{2t}} \mathrm{d}t = \frac{1}{\sqrt{2\alpha}} \mathrm{e}^{-\sqrt{2\alpha}|x|}, \ x \in \mathbf{R}.$$

6. ($[0,1]$ 上反射 Brown 运动)设 ψ 是周期 2 的函数, $\psi(x) := 1 - |x-1|$, $x \in [0,2]$. 证明: $\psi(B.)$ 是 $[0,1]$ 上的 Markov 过程. 求其生成算子.
7. 为了避免一个等价类的麻烦, 我们通常在一个比有界可测函数更小的函数类上来考虑生成算子. 引入函数类

$$\mathscr{D}(E) := \{f \in \mathscr{E}_b : \text{对任何 } x \in E, \ f(X_t) \longrightarrow f(X_0) \text{ ess. } \mathbb{P}^x\},$$

其中 ess. $\mathbb{P}^x$ 指对任何 $t_n \downarrow 0$, $\lim_n f(X_{t_n}) = f(X_0)$ a.s. $\mathbb{P}^x$. 证明:

(1) $C_u(E) \subset \mathscr{D}(E)$;

(2) $U^\alpha \mathscr{E}_b \subset \mathscr{D}(E)$, 因此 $U^\alpha \mathscr{D}(E) \subset \mathscr{D}(E)$;

(3) 对任何 $f \in \mathscr{D}(E)$, 过程 $f(X.)$ 右连续;

(4) $U^\alpha \mathscr{D}(E)$ 与 α 无关;

(5) 如果 $f \in \mathscr{D}(E)$, $U^\alpha f = 0$, 则 $f = 0$.

然后定义 $D(A) = U^\alpha \mathscr{D}(E)$, $Au = \alpha u - f$. 如果 $u = U^\alpha f$, $f \in \mathscr{D}(E)$, 那么算子 $(A, D(A))$ 也唯一地决定 $\{\mathbb{P}^x\}$.

8. 设 $\{\mathbb{P}^x : x \in E\}$ 是一个 Markov 过程. G 是 E 的开子集, T 是 G 的首离时. 定义 $Q_t(x, A) := \mathbb{P}^x(X_t \in A, t < T)$. 证明: (Q_t) 是 $(E, \mathscr{E}^*)$ 上的转移函数半群. (它所对应的 Markov 过程称为是 $\{\mathbb{P}^x\}$的子过程, 这个变换称为是 killing 变换.)

§4.2 过分函数与精细拓扑

继续讨论右过程 X 的性质. 注意半群 (P_t) 可以自动地看成 $(E,\mathscr{E}^*)$ 上的半群. 对 $\alpha \geqslant 0$, 记 $P_t^\alpha := \mathrm{e}^{-\alpha t}P_t$ 也是一个转移半群. 设 f 是一个取值 $[0,\infty]$ 的普遍可测函数. 如果对任何 $t>0$, $P_t^\alpha f \leqslant f$, 说 f 是 α-上平均的. 如果 f 是 α-上平均的, 那么 $t \mapsto P_t^\alpha f$ 递减.

引理 4.2.1 设 f 是非负普遍可测函数且 $f(X_\cdot)$ 可积, 则 f 是 α-上平均的当且仅当对任何 $x \in E$, 过程 $\{\mathrm{e}^{-\alpha t}f(X_t)\}$ 是 $(\Omega,\mathscr{F},\mathbb{P}^x)$ 上 $(\mathscr{F}_t)$ 上鞅.

证明 设 f 是 α-上平均的, 由 Markov 性,

$$\begin{aligned}\mathbb{E}^x(\mathrm{e}^{-\alpha(t+s)}f(X_{t+s})|\mathscr{F}_t) &= \mathrm{e}^{-\alpha(t+s)}\mathbb{E}^{X_t}f(X_s)\\ &= \mathrm{e}^{-\alpha(t+s)}P_sf(X_t) \leqslant \mathrm{e}^{-\alpha t}f(X_t).\end{aligned}$$

因此按定义, 过程 $\{\mathrm{e}^{-\alpha t}f(X_t),\mathscr{F}_t\}$ 是上鞅. 反过来, 由上鞅定义,

$$P_t^\alpha f(x) = \mathbb{E}^x(\mathbb{E}^x(\mathrm{e}^{-\alpha t}f(X_t)|\mathscr{F}_0)) \leqslant \mathbb{E}^x f(X_0) = f(x),$$

故 f 是 α-上平均的. □

函数 f 称为是 α-过分的, 如果它是 α-上平均的且

$$f = \uparrow \lim_{t\downarrow 0} P_t^\alpha f.$$

用 $\mathbf{S}^\alpha(X)$ 或 $\mathbf{S}^\alpha$ 表示 α-过分函数全体, 当 $\alpha = 0$ 时, 就省略这个符号. 设 f 是 α-上平均的, 令 $\hat{f} :=\uparrow \lim_{t\downarrow 0} P_t^\alpha f$, 称它是 f 的正则化. 显然 $f \geqslant \hat{f}$. 由单调收敛定理不难验证下列结论.

定理 4.2.1 设 f 是 α-上平均的, 则 $\hat{f} \in \mathbf{S}^\alpha$ 且是被 f 控制的最大 α-过分函数.

我们现列出 $\mathbf{S}^\alpha$ 的一些简单性质.

引理 4.2.2 过分函数有下列性质:

(1) $\mathbf{S}^\alpha$ 是一个锥;

(2) $\mathbf{S}^\alpha$ 对递增函数列极限封闭;

(3) 如果 $\alpha > r \geqslant 0$, 那么 $\mathbf{S}^\alpha \supset \mathbf{S}^r$ 且 $\mathbf{S}^r = \bigcap_{\alpha > r} \mathbf{S}^\alpha$;

(4) 如果 f 是非负 $\mathscr{E}^*$-可测的, 那么 f 的 α-位势 $U^\alpha f \in \mathbf{S}^\alpha$;

(5) $f \in \mathbf{S}^\alpha$ 当且仅当 $\beta \uparrow +\infty$ 时, $\beta U^{\alpha+\beta} f$ 递增收敛于 f;

(6) 设 $\alpha > 0$, $f \in \mathbf{S}^\alpha$, 则存在 $g_n \in b\mathscr{E}_+^*$ 使得 $U^\alpha g_n \uparrow f$.

证明 (1), (2) 及 (3) 的前半部分由定义是显然的. 下面验证 (3) 的第二部分. 包含关系 $\subset$是显然的, 我们来验证反包含. 设 $f \in \bigcap_{\alpha > r} \mathbf{S}^\alpha$, 则映射 $\alpha \mapsto P_t^\alpha f(x)$ 与 $t \mapsto P_t^\alpha f(x)$ 是递减的, $P_t^r f = \lim_{\alpha \downarrow r} P_t^\alpha f \leqslant f$ 且

$$\uparrow \lim_{t \to 0} P_t^r f = \uparrow \lim_{t \to 0} \uparrow \lim_{\alpha \downarrow r} P_t^\alpha f = \uparrow \lim_{\alpha \downarrow r} \uparrow \lim_{t \to 0} P_t^\alpha f = f.$$

因此 $f \in \mathbf{S}^r$.

(4) 当 $t \downarrow 0$, 由半群性质,

$$\mathrm{e}^{-\alpha t} P_t U^\alpha f(x) = \int_t^\infty \mathrm{e}^{-\alpha s} P_s f(x) \mathrm{d}s \uparrow U^\alpha f(x).$$

因此 $U^\alpha f \in \mathbf{S}^\alpha$.

对于 (5), 验证充分性. 先设 f 有界, 则当 $\beta > 0$, $U^{\beta+\alpha} f$ 有界. 由豫解方程 $U^{\beta+\alpha} f = U^\alpha (f - \beta U^{\beta+\alpha} f) \in \mathbf{S}^\alpha$. 一般地, 由单调收敛定理, $U^{\beta+\alpha}(n \wedge f) \uparrow U^{\beta+\alpha} f$, 故 $U^{\beta+\alpha} f \in \mathbf{S}^\alpha$. 因此 f 是 $\mathbf{S}^\alpha$ 中递增函数列 $nU^{n+\alpha} f$ 的极限, 由上面性质 (2), $f \in \mathbf{S}^\alpha$. 反之, 如果 $f \in \mathbf{S}^\alpha$, 那么对任何 $\beta > 0$,

$$\lim_{\beta \to \infty} \beta U^{\beta+\alpha} f = \lim_{\beta \to \infty} \int_0^\infty \mathrm{e}^{-t} P_{\frac{t}{\beta}}^\alpha f \mathrm{d}t = f.$$

因此 $\beta \mapsto \beta U^{\alpha+\beta} f$ 递增且由单调收敛定理, $\lim_{\beta \to \infty} \beta U^{\alpha+\beta} f = f$.

最后证明 (6). 由 (5) 得 $nU^{n+\alpha} f \uparrow f$, 令

$$g_{n,k} := n(f \wedge k - nU^{n+\alpha}(f \wedge k)),$$

则 $g_{n,k} \in b\mathscr{E}_+^*$ 且 $U^\alpha g_{n,k}$ 关于 n, k 都递增且收敛于 f. 令 $g_n := g_{n,n}$, 则 $U^\alpha g_n \uparrow f$. □

由 X 的简单 Markov 性不难验证: 如果 f 是有界非负可测函数, $\alpha > 0$, 那么上鞅 $(\mathrm{e}^{-\alpha t} U^\alpha f(X_t))$ 有下面的 Doob-Meyer 分解:

$$\mathbb{E}^x\left(\int_0^\infty \mathrm{e}^{-\alpha s}f(X_s)\mathrm{d}s|\mathscr{F}_t\right)=\int_0^t \mathrm{e}^{-\alpha s}f(X_s)\mathrm{d}s+\mathrm{e}^{-\alpha t}U^\alpha f(X_t).$$

引理 4.2.2 的 (6) 说明, 当 $\alpha>0$ 时, α-过分函数是 α-位势函数的递增极限. 此结论当 $\alpha=0$ 时一般是不成立的. 为了看清这个问题, 我们先来看所谓的位势核, 直观地讲, 对 $x\in E$, $A\in\mathscr{E}$, 位势核

$$U(x,A):=U^0(x,A)=\mathbb{E}^x\int_0^\infty 1_A(X_t)\mathrm{d}t,$$

即 $U(x,A)$ 是过程从 x 点出发, 逗留在集合 A 中的平均时间. 一般地, $U(x,\cdot)$ 是一个 s-有限测度, 未必是 σ- 有限的, 更不同于 α-位势核 $U^\alpha(x,\cdot)$ 是有限的.

定义 4.2.1　半群 (P_t) 称为暂留的, 如果存在集列 $K_n\uparrow E$, $K_n\in\mathscr{E}^*$, 使得 $U(\cdot,K_n)$ 是有界的.

显然如果 $\alpha>0$, 那么半群 (P_t^α) 是暂留的. 暂留的直观意义是过程的轨道处在一个有界区域内的平均时间是有限的.

例 4.2.1　计算 $\mathbf{R}^d$ 上热核半群的位势核, 则

$$u(x):=\int_0^\infty \frac{1}{(2\pi t)^{\frac{d}{2}}}\mathrm{e}^{-\frac{|x|^2}{2t}}\mathrm{d}t=\int_0^\infty \frac{t^{\frac{d}{2}-2}}{2(\pi)^{\frac{d}{2}}}\mathrm{e}^{-t\frac{|x|^2}{2}}\mathrm{d}t.$$

因此当 $d=1,2$ 时, $u\equiv+\infty$. 当 $d\geqslant 3$ 时,

$$u(x)=\frac{\Gamma(\dfrac{d}{2}-1)}{2\pi^{\frac{d}{2}}|x|^{d-2}}.$$

而位势核 $U(x,\mathrm{d}y)=u(y-x)\mathrm{d}y$. 容易验证当 $d=1,2$ 时, $U1_A\equiv 0$ 或 ∞. 这时热核半群是常返的, 其过分函数不可能是位势的极限. 当 $d\geqslant 3$ 时, 热核半群是暂留的. 这说明 3-维或 3-维以上的 Brown 运动最终将走向无穷远处.

对于暂留的半群, 过分函数仍然是位势函数的递增极限.

定理 4.2.2　设半群 (P_t) 是暂留的, $f\in\mathbf{S}$, 则存在 $g_n\in b\mathscr{E}_+$, 使得 $Ug_n\uparrow f$ 且对每个 n, g_n 与 Ug_n 是有界的.

证明　取以上定义中所言的集列 $\{K_n\}$, 令 $h_n:=n1_{K_n}$, 则 Uh_n 有界且 $h_n\uparrow+\infty$.

再令 $f_n := Uh_n \wedge f$, 那么 f_n 是上平均的, 有界且递增收敛于 f. 当 $t \uparrow +\infty$ 时, $P_t f_n \leqslant P_t U h_n = \int_t^\infty P_s h_n \mathrm{d}s \downarrow 0$. 令

$$g_{n,k} := k(f_n - P_{\frac{1}{k}} f_n),$$

则

$$\begin{aligned}\int_0^t P_s g_{n,k} \mathrm{d}s &= k\left(\int_0^t P_s f_n \mathrm{d}s - \int_{\frac{1}{k}}^{t+\frac{1}{k}} P_s f_n \mathrm{d}s\right)\\ &= k\left(\int_0^{\frac{1}{k}} P_s f_n \mathrm{d}s - \int_t^{t+\frac{1}{k}} P_s f_n \mathrm{d}s\right).\end{aligned}$$

因此 $Ug_{n,k} = k\int_0^{\frac{1}{k}} P_s f_n \mathrm{d}s$, 它关于 n, k 递增收敛于 f. 故若 $g_n = g_{n,n}$, 则 $Ug_n \uparrow f$ 且 $g_n \leqslant nf_n$, $Ug_n \leqslant f_n$ 有界. □

定理 4.2.3　对任何 $\alpha > 0$, $f \in \mathbf{S}^\alpha$, 过程 $f(X.)$ 是右连续且存在左极限的.

证明　存在 $g_n \geqslant 0$, 使得 $U^\alpha g_n \uparrow f$. 因为 $\mathrm{e}^{-\alpha t} U^\alpha g_n(X_t)$ 是右连续上鞅, 故 $U^\alpha g_n(X.)$ 是右连续且存在左极限的. 而 $\mathrm{e}^{-\alpha t} U^\alpha g_n(X_t) \uparrow \mathrm{e}^{-\alpha t} f(X_t)$, 因此由 Meyer 定理(参考 §3.3 习题或文献[15]), 过程 $f(X.)$ 是右连续且左极限存在. □

E 上的非负函数 f 称为是近乎 Borel 可测的, 如果对 E 上任何概率 μ, 存在 Borel 可测函数 f_1, f_2, 使得 $f_1 \leqslant f \leqslant f_2$ 且过程 $(f_1(X_t))$ 与 $(f_2(X_t))$ 是 $\mathbb{P}^\mu$ 不可区分的, 即 X 不能将其与一个 Borel 可测函数区分开. 我们用 $\mathscr{E}^n$ 表示 E 上近乎 Borel 可测集全体, 则 $\mathscr{E}^n$ 是 σ- 代数且那么 $\mathscr{E} \subset \mathscr{E}^n \subset \mathscr{E}^*$. 一个 $\mathscr{E}^*$ 可测函数 f 称为是可选的, 如果 $f(X.)$ 是(不可区分于)一个 $(\mathscr{F}_t)$ 可选过程; 称为是近乎可选的, 如果对任何 E 上概率 μ, $f(X.)$ 是($\mathbb{P}^\mu$-不可区分于)一个 $(\mathscr{F}_t^\mu)$-可选过程. 子集 $B \in \mathscr{E}^*$ 称为可选或近乎可选, 如果其指标 1_B 是. 容易验证近乎 Borel 集一定是近乎可选的. 另外由定理 4.2.3 推出 α-过分函数是可选的.

定理 4.2.4　设 $\alpha > 0$. 如果 X 是 Borel 右过程, 则任何 α-过分函数是近乎 Borel 可测的.

证明　不妨设 $f = U^\alpha h$, $h \in b\mathscr{E}_+^*$. 对 E 上的概率测度 μ, 存在 $h_1, h_2 \in b\mathscr{E}_+$, 使得 $h_1 \leqslant h \leqslant h_2$ 且 $\mu U^\alpha(\{h_1 \neq h_2\}) = 0$, 那么 $U^\alpha h_1 \leqslant f \leqslant U^\alpha h_2$ 且对任何 $t \geqslant 0$,

$$\mathbb{E}^{\mu}(U^{\alpha}(h_2-h_1)(X_t))=\mathrm{e}^{\alpha t}\mu P_t^{\alpha}U^{\alpha}(h_2-h_1)\leqslant \mathrm{e}^{\alpha t}\mu U^{\alpha}(h_2-h_1)=0.$$

因此对任何 $t\geqslant 0$, $U^{\alpha}h_2(X_t)=U^{\alpha}h_1(X_t)$, $\mathbb{P}^{\mu}$-a.s.. 由右连续性推出两个过程是 $\mathbb{P}^{\mu}$-不可区分的. 因此 $f\in\mathscr{E}^n$. □

用 $\mathscr{E}^e$ 表示由 $\bigcup_{\alpha\geqslant 0}\mathbf{S}^{\alpha}$ 生成的 σ- 代数, 定理说明如果 X 是 Borel 右过程, 那么 $\mathscr{E}\subset\mathscr{E}^e\subset\mathscr{E}^n\subset\mathscr{E}^*$. 一般右过程不一定对.

定理 4.2.5　设 X 是右过程, 则

(1) P_t 是 $(E,\mathscr{E}^e)$ 上的核;

(2) $\mathscr{E}^e$ 可测函数是可选的;

(3) $\mathscr{E}\subset\mathscr{E}^e$.

证明　(1) 成立是因为 $P_t\mathbf{S}^{\alpha}\subset\mathbf{S}^{\alpha}$.

(2) 若 $f\in\mathbf{S}^{\alpha}$, 则由定理 4.2.3, 过程 $f(X_\cdot)$ 是右连续且有左极限, 故过程 $(f\circ X_t)$ 是可选的. 然后由单调类定理推出 (2) 成立.

(3) 对任何 $f\in C_u(E)$, 由 X 的右连续性, 当 $\alpha\to\infty$ 时, $\alpha U^{\alpha}f$ 点点收敛于 f, 因此 $f\in\mathscr{E}^e$, 故 $\mathscr{E}\subset\mathscr{E}^e$. □

下面所谓停时都是指 $(\mathscr{F}_t)$-停时. 实现如果 $A\in\mathscr{E}$, 那么由定理 3.1.6 看出对任何 $t>0$, $\{D_A<t\}\in(\mathscr{F}_t^0)^*\subset\mathscr{F}_t=\mathscr{F}_{t+}$, 因此 D_A 和 T_A 是停时. 它们是停时这个结论是我们强化自然流的最主要的理由之一. 但我们有更一般的结论.

定理 4.2.6　如果 A 是近乎可选的, 那么 A 的进入时与首中时 D_A, T_A 是停时.

证明　由定义知, 对任何 E 上概率 μ, $1_A(X_\cdot)$ 是 $(\mathscr{F}_t^{\mu})$-可选的, 因此是 $(\mathscr{F}_t^{\mu})$-循序可测的. 由定理 3.1.7 而 X 对于 A 的进入时或首中时就是 $1_A(X)$ 对于 Borel 集 $\{1\}$ 的进入时或首中时. 故由定理 3.1.7, D_A, T_A 是 $(\mathscr{F}_t^{\mu})$-停时, 因为 μ 是任意的, 故它们是 $(\mathscr{F}_t)$ 停时. □

现在我们引入正则点的概念.

定义 4.2.2　设 A 是近乎可选集, $x\in E$. 称 x 关于 A 是正则的 (或 A 的正则点), 如果

$$\mathbb{P}^x(T_A=0)=1.$$

用 A^r 表示 A 的正则点全体, 记 $\tilde{A}:=A\cup A^r$, 并称它是 A 的精细闭包.

因为 $\{T_A=0\}\in\mathscr{F}_0$, 由 Blumenthal 0-1 律, $x\notin A^r$ 当且仅当 $\mathbb{P}^x(T_A>0)=1$.

直观地, $x \in A^r$ 意味着从 x 点出发的轨道将立刻遇到 A. 要注意的是 x 可以同时是 A 及其补集 A^c 的正则点. 如果 A 是闭集, 则由轨道右连续性, 从开集 A^c 内出发的过程必将在 A^c 中滞留若干时间, 因此 A^c 中的点不可能是 A 的正则点, 即 $A^r \subset A$. 令 $\phi_B^1(x) := \mathbb{E}^x(\mathrm{e}^{-T_B})$, $x \in E$, 那么容易验证 ϕ_B^1 是 1-过分函数. 因此 $B^r = \{\phi_B^1 = 1\} \in \mathscr{E}^e$.

定理 4.2.7 设 A 是近乎可选的, 则在 $\{T_A < \infty\}$ 上, $X_{T_A} \in \tilde{A}$ a.s. 换句话说, 对任何 $x \in E$,

$$\mathbb{P}^x(X_{T_A} \notin \tilde{A}, T_A < \infty) = 0.$$

证明 由 T_A 的定义不难看出, 当 $T_A < \infty$ 时, $\{X_{T_A} \notin A\} \subset \{T_A \circ \theta_{T_A} = 0\}$, 因此由强 Markov 性得

$$\begin{aligned}
&\mathbb{P}^x(X_{T_A} \notin \tilde{A}, T_A < \infty)\\
&= \mathbb{P}^x(X_{T_A} \notin A, X_{T_A} \notin A^r, T_A < \infty)\\
&\leqslant \mathbb{P}^x(X_{T_A} \notin A^r, T_A < \infty, T_A \circ \theta_{T_A} = 0)\\
&= \mathbb{E}^x(X_{T_A} \notin A^r, T_A < \infty, \mathbb{P}^{X(T_A)}(T_A = 0)) = 0.
\end{aligned}$$

完成证明. □

一个随机时间 T 称为原始的终止时, 如果对任何 $t \geqslant 0$, $T > t$ 蕴含着 $T = t + T\circ\theta_t$; 如果它还是停时, 那么称为终止时. 如果 T 是原始的终止时, 则 $t \mapsto t + T\circ\theta_t$ 是递增的, 它称为是恰好的, 如果 $T = \downarrow \lim_{t\to 0}(t + T\circ\theta_t)$. 如果 T 是终止时, 则 $t + T\circ\theta_t \geqslant T \vee t$. 一个集合 $B \in \mathscr{E}^e$ 的首中时与进入时都是终止时, 首中时是恰好的, 但进入时一般不是恰好的. 定义

$$P_T^\alpha(x, A) = \mathbb{E}^x(\mathrm{e}^{-\alpha T}, X_T \in A, T < \infty).$$

显然 P_T^α 是 $(E, \mathscr{E}^*)$ 上的核. 如果 T 是集合 B 的首中时, 那么写 P_T^α 为 P_B^α. 另外定义 $\Phi_B^\alpha := P_B^\alpha 1 = \mathbb{E}^x \mathrm{e}^{-\alpha T_B}$. 定理 4.2.7 蕴含着测度 $P_B^\alpha(x, \cdot)$ 支撑在 $\tilde{B}$ 上, 且当 $x \in B^r$ 时, 它是 x 点的 Dirac 测度. 也称 P_B^α 为扫除算子, 因为它将 x 点上的 Dirac 测度"扫"为 $\tilde{B}$ 上的测度. 扫除算子是位势理论中的重要概念, 它的引入要归功于 Poincaré, 但 Hunt 建立了其概率表达式.

定理 4.2.8 (1) 如果 T 是终止时, $\alpha \geqslant 0$, $f \in \mathbf{S}^\alpha$, 则 $P_T^\alpha f$ 是 α-上平均的. 进一步, 如果 T 是恰好的, 则 $P_T^\alpha f \in \mathbf{S}^\alpha$.

(2) 如果 T 是停时, $\alpha \geqslant 0$, $f \in \mathbf{S}^\alpha$, 那么 $P_T^\alpha f \leqslant f$.

证明 (1) 先考虑位势, 由 Dynkin 公式,

$$P_t^\alpha P_T^\alpha U^\alpha f(x) = \mathbb{E}^x \int_{t+T\circ\theta_t}^\infty \mathrm{e}^{-\alpha s} f(X_s)\mathrm{d}s \leqslant P_T^\alpha U^\alpha f.$$

显然如果 T 是恰好的, 则 $P_T^\alpha U^\alpha f \in \mathbf{S}^\alpha$. 当 $\alpha > 0$ 时, 对 $f \in \mathbf{S}^\alpha$, 存在 $U^\alpha f_n \uparrow f$, 由单调收敛定理推出 $P_T^\alpha f \in \mathbf{S}^\alpha$. 最后对 $f \in \mathbf{S}$, 则当 $r > \alpha > 0$ 时, $P_T^\alpha f \in \mathbf{S}^r$. 让 $\alpha \downarrow 0$, 则 $P_T^\alpha f \uparrow P_T f$, 因此 $P_T f \in \mathbf{S}^r$, 而 r 任意, 故 $P_T f \in \mathbf{S}$.

(2) 先设 $\alpha > 0$. 也先取位势 $f = U^\alpha g$, $g \in \mathscr{E}_+^*$. 由 Dynkin 公式,

$$P_T^\alpha U^\alpha g = \mathbb{E}^x \int_T^\infty \mathrm{e}^{-\alpha t} g(X_t)\mathrm{d}t \leqslant U^\alpha g(x).$$

类似地, 由单调收敛定理推出对 $f \in \mathbf{S}^{alpha}$, 有 $P_T^\alpha f \leqslant f$. 若 $\alpha = 0$, 则对任何 $r > 0$, $P_T^r f \leqslant f$. 因此 $P_T f \leqslant f$. □

例 4.2.2 设 B 是 $\mathbf{R}^d$ 上 Brown 运动,

$$A = \{x \in \mathbf{R}^d:\ |x| > 1\}.$$

显然若 $|x| \geqslant 1$, 则 x 是 A 的正则点, 故 $P_A(x,\cdot) = \epsilon_x$; 若 $|x| < 1$, 则 $\mathbb{P}^x(X_{T_A} \in \partial A) = 1$, 即 X_{T_A} 位于单位球面上且 P_A 是 Poisson 核, 即

$$P_A f(x) = \int_{\{y:|y|=1\}} \frac{1-|x|^2}{|x-y|^2} f(y)\sigma(\mathrm{d}y),$$

其中 σ 是单位球面上的均匀分布. ▌

下面我们将引入 E 上的所谓精细拓扑, 并证明它是使得 $\mathbf{S}^\alpha$ 中函数成为连续函数的最小拓扑.

定义 4.2.3 集合 $A \subset E$ 称为是一个精细开集, 如果对任何 $x \in A$, 存在近乎可选集 B, 使得 $x \in B \subset A$ 且 $\mathbb{P}^x(T_{B^c} > 0) = 1$.

注 定义中的 B 还可以被要求是紧的. 事实上, 存在包含 $\varDelta$ 的递减开集列 $\{G_n\}$, 使得 $G_n \supset B^c$ 且

$$\mathbb{P}^x(\uparrow \lim_n T_{G_n} = T_{E_\Delta \setminus B}) = 1.$$

由 Blumenthal 0-1 律, 存在 n, 使得 $\mathbb{P}^x(T_{G_n} > 0) = 1$. 令 $K := G_n^c$, 则 K 紧 $K \subset B$ 且 $\mathbb{P}^x(T_{K^c} > 0) = 1$.

也就是说, A 是精细开集当且仅当从 A 中点出发的任何轨道都会在 A 中滞留一段时间. 由右连续性, 任何开集是精细开集. 用 $\mathscr{O}$ 表示 E 上的精细开集全体, 那么容易验证 $\mathscr{O}$ 是 E 上的拓扑, 称为是 E 上的精细拓扑, 它比 E 上原来的拓扑更细. 一个函数称为精细连续, 如果它在精细拓扑下连续.

例 4.2.3 设 X 是 $\mathbf{R}$ 上向右一致平移, 那么左闭右开区间 $[a, b)$ 是精细开集. 另外对任何 $x \in \mathbf{R}$, 如果 A 是包含 x 的一个精细开集, 那么存在 $\delta > 0$ 使得 $[x, x+\delta) \subset A$, 也就是说, x 有一个包含于 A 中的左闭右开的邻域. 因此左闭右开区间全体是精细拓扑的一个基, 故一个函数精细连续当且仅当它是右连续的. ▮

定理 4.2.9 设 $\alpha > 0$, 则 E 上的精细拓扑是使得 $\mathbf{S}^\alpha$ 中函数成为连续函数的最小拓扑. 换句话说, 精细拓扑是 $\mathbf{S}^\alpha$ 生成的拓扑.

证明 首先证明 α-过分函数 f 是精细拓扑下连续的. 实际上只要 $f(X_t)$ 在 $t=0$ 处右连续, 那么对开区间 I, 令 $B := f^{-1}(I)$, 则 $B \in \mathscr{E}^e$ 且 $B^c = f^{-1}(I^c)$. 对任何 $x \in B$, $f(x) \in I$. 由右连续性, $\mathbb{P}^x(T_{B^c} = 0) \leqslant \mathbb{P}^x(\lim_{t\to 0} f(X_t) \in I^c) = 0$. 因此 $B \in \mathscr{O}$. 故 f 精细连续.

设 $\mathscr{K}$ 是 E 的紧子集全体, 令

$$\mathscr{U} := \{\{\varPhi^\alpha_{K^c} < 1\} : K \in \mathscr{K}\}.$$

因为 $\varPhi^\alpha_{K^c} \in \mathbf{S}^\alpha$ 是精细连续的, 故 $\mathscr{U} \subset \mathscr{O}$. 只要我们能验证 $\mathscr{U}$ 是 $\mathscr{O}$ 的一个拓扑基就完成了定理的证明. 事实上, 设 A 是点 x 的一个精细邻域, 由上面定义下的注推出存在 $K \in \mathscr{K}$, 使得 $x \in K \subset A$ 且 $\mathbb{P}^x(T_{K^c} > 0) = 1$. 令 $B := \{\varPhi^\alpha_{K^c} < 1\}$, 则 $B \in \mathscr{U}$ 且因为 K^c 是开集, 故 $K^c \subset B^c$, 即 $K \supset B$. 另外因 $x \in B$, 故 B 也是 x 的精细邻域且 $B \subset A$. □

现在我们给出精细连续的一个刻画.

定理 4.2.10 设 f 是 E 上近乎可选函数, 则 f 是精细连续的当且仅当 $f(X_\cdot)$ 是右连续的.

证明 实际上, 由 $t \mapsto f(X_t)$ 在 $t=0$ 处的右连续性即可推出 f 是精细连续的. 反过来, 设 f 是精细连续的. 我们只要证明 $f(X_\cdot)$ 的几乎所有轨道右下半连续就够了, 因为同样方法可以证明它的几乎所有轨道是右上半连续的. 用 N 表示 $f(X_\cdot)$ 的不是右下半连续的轨道全体, 如果 $\omega \in N$ 且 $t \mapsto f\circ X_t(\omega)$ 在某点 $t_0 \geqslant 0$ 处不是右下半连续的, 即 $\liminf_{t\downarrow t_0} f(X_t(\omega)) < f(X_{t_0}(\omega))$, 那么存在 $r \in \mathbb{Q}$ 与 $t_n \downarrow\downarrow t_0$, 使得 $X_{t_n}(\omega) \in \{f<r\}$, $X_{t_0}(\omega) \in \{f>r\}$. 令 $B_r := \{f<r\}$, $\phi_r := \Phi^1_{B_r}$, 那么 $B_r \in \mathscr{O}$, $\phi_r(X_{t_n}(\omega)) = 1$ 且 $\phi_r(X_{t_0}(\omega)) < 1$. 因此 $N \subset \bigcup_{r\in\mathbb{Q}}\{\omega\in\Omega : t\mapsto \phi_r(X_t(\omega))$ 不右连续$\}$ 内, 而 $\phi_r \in \mathbf{S}^1$, 因此右边是零概率事件的可列并, 从而对任何 $x \in E$, $\mathbb{P}^x(N)=0$, 故过程 $f(X_\cdot)$ 是右下半连续的. □

状态空间上以过程 X 的角度看有几种小集合的概念.

定义 4.2.4 给定右过程 X.

(1) 集合 $A \in \mathscr{E}^*$ 称为是位势零集, 如果 $U(\cdot,A)\equiv 0$;

(2) 集合 $A \subset E$ 称为是极集, 如果存在近乎可选集 B, 使得 $A \subset B$ 且对任何 $x \in E$, $\mathbb{P}^x(T_B<\infty)=0$;

(3) 集合 $A \subset E$ 称为是瘦的, 如果存在近乎可选集 B 使得 $A\subset B$ 且 $B^r=\varnothing$;

(4) 集合 $A \subset E$ 称为半极集, 如果它包含于一瘦集列的并中.

注 因为 $U(x,A) = \mathbb{E}^x\int_0^\infty 1_A(X_t)\mathrm{d}t = \mathbb{E}^x(|\{t : X_t\in A\}|)$, 故直观地, 位势 0 集是过程滞留在其上的平均时间为零的集合, 它不能含有精细内点, 且 $U(\cdot,A)\equiv 0$ 等价于对某个(因此对所有) $\alpha \geqslant 0$, $U^\alpha(\cdot,A)\equiv 0$. 极集是过程从任何点出发都不能到达的集合, 且 $A\in\mathscr{E}^e$ 是极集当且仅当对某个(因此对所有) $\alpha\geqslant 0$, $\Phi^\alpha_A\equiv 0$. 瘦集是过程从任何点出发都不会滞留的集合, 且 $A\in\mathscr{E}^e$ 是瘦集当且仅当对某个(因此对所有) $\alpha>0$, $\Phi^\alpha_A<1$. 显然极集是瘦的, 而瘦集是半极集.

例 4.2.4 设 X 是 $\mathbf{R}$ 上向右一致平移, 那么仅有的极集是空集, 单点集是瘦集, 任何可数集是半极集, 任何 Lebesgue 零测集是位势零集.

定理 4.2.11 设 A 是半极集, 则几乎肯定有 $\{t\geqslant 0 : X_t\in A\}$ 是可数的. 因此半极集是位势零集.

证明 不妨设 A 是瘦的且是 Borel 集. 令 $B := A\cap\{\Phi^1_A\leqslant a\}$, $a<1$. 因为 A 是形如 B 的集合的可列并, 因此只需证明几乎肯定有仅有可列次 $X_t\in B$ 就够了. 令 $T_1 := T_B$, $T_{n+1} := T_n + T_1\circ\theta_{T_n}$. T_n 是第 n 到达 B 的时间. 因为 B^r 空, 故由定理

4.2.7, 当 $T_n<\infty$ 时, $X_{T_n}\in B$. 用强 Markov 性,

$$\mathbb{E}^x\mathrm{e}^{-T_{n+1}}=\mathbb{E}^x(\mathrm{e}^{-T_n}\mathbb{E}^{X(T_n)}\mathrm{e}^{-T_1})\leqslant a\mathbb{E}^x\mathrm{e}^{-T_n},$$

那么 $\mathbb{E}^x\mathrm{e}^{-T_n}\to 0$. 由有界收敛定理 $\mathbb{P}^x(T_n\to\infty)=1$, $x\in E$. 由此推出定理的结论. □

定理 4.2.12 设 A 是近乎可选的, 则 $A\setminus A^r$ 是半极集.

证明 令 $A_n:=A\cap\{\Phi_A^1\leqslant 1-\dfrac{1}{n}\}$, 那么 $A\setminus A^r=\bigcup\limits_n A_n$ 且 A_n 是瘦的. 事实上, 如果 $\Phi_A^1(x)<1$, 那么 x 不是 A 的也不是 A_n 的正则点. 如果 $\Phi_A^1(x)=1$, 因为 Φ_A^1 是精细连续的, 故 x 所在的集合 $\{\Phi_A^1>1-\dfrac{1}{n}\}$ 是精细开集. 而它与 A_n 不相交, 推出 x 不能是 A_n 的正则点. 因此 A_n 没有正则点, 是瘦集. □

定理 4.2.13 (1) 设 f 是 α-上平均的, $\overline{f}$ 是由 f 控制的最大 α-过分函数, 则 $\{f>\overline{f}\}$ 是位势零集;

(2) 设 $f,g\in\mathbf{S}^\alpha$, 如果 $f\geqslant g$ 在一个位势 0 集外成立, 那么它处处成立.

证明 (1) 因为 $\overline{f}=\uparrow\lim_{t\downarrow 0}P_t^\alpha f$, 所以推出 $U^1f=U^1\overline{f}$, 因此 $\{f>\overline{f}\}$ 是位势 0 集. (2) 由条件推出对任何 n, $nU^{n+\alpha}f\geqslant nU^{n+\alpha}g$ 恒成立, 令 n 趋于无穷得 $f\geqslant g$ 恒成立. □

习 题

1. 集合 $A\in\mathscr{E}^*$ 称为是吸收的, 如果对任何 $x\in A$, $t>0$, $\mathbb{P}^x(X_t\in A^c)=0$. 证明: 若 f 是过分函数, 那么 $\{f<\infty\}$ 是吸收集.
2. 对任何 $x\in\mathbf{R}$, $t>0$, 定义 $\mathbb{P}^x(X_t=x+t)=1$. $\{\mathbb{P}^x\}$ 称为是右一致平移. 试刻画右一致平移过程的上平均函数类, 过分函数类及其精细拓扑.
3. 证明: Brown 运动或热核半群的过分函数是下半连续的.
4. 证明: 过程暂留当且仅当存在 $B_n\in\mathscr{E}^e$, $B_n\uparrow E$, 使得对任何 n, $x\in E$, 有 $\mathbb{P}^x(L_{B_n}<\infty)=1$, 其中 L_B 表示末离时 $L_B:=\sup\{t:X_t\in B\}$.
5. 证明: 极集的可列并还是极集, 半极集的可列并还是半极集.
6. 设 A 是位势零集. 证明: E 的所有点都是 A 的正则点.
7. 设 $A\in\mathscr{E}$, 且 $\limsup_{t\downarrow 0}P_t(x,A)>0$. 证明: x 是 A 的正则点.

8. 如果单点集总是极集, 证明: 序列 $\{x_n\}$ 在精细拓扑下收敛于 x 当且仅当对充分大的 n 有 $x_n = x$. (在这种情况下, 精细拓扑不满足第一可数公理.) 再证明: Brown 运动的精细拓扑不是离散的.
9. 证明: 精细拓扑空间是完全正则 Hausdorff 空间.
10. 如果 B 是位势零集. 证明: $E\backslash B$ 在 E 中精细稠密.
11. 设对任何非空精细开集 B 及 $x \in E$ 有 $\mathbb{P}^x(T_B < \infty) = 1$. 证明: 任何过分函数是常数.

§4.3 Feller 过程与 Lévy 过程

在这一节, 我们证明满足 Feller 性的半群有一个 Borel 右过程的实现并介绍 Lévy 过程. 设 E 是一个局部紧具有可数基的 Hausdorff 空间, 加一点 Δ, 用 $E_\Delta := E \cup \{\Delta\}$ 表示 E 的 Alexander 紧化. $C_0(E)$ 表示无穷远 Δ 处为零的连续函数全体, $C_c(E)$ 表示 E 上具紧支撑连续函数全体, 则 $C_c(E) \subset C_0(E) \subset C(E_\Delta) \subset C_b(E)$. 用 $\|\cdot\|$ 表示一致范数, 则 $C(E_\Delta)$ 是 Banach 空间, 而 $C_0(E)$ 是其一闭子空间. 实际上, 对任何 $f \in C(E_\Delta)$, $f - f(\Delta)1_{E_\Delta} \in C_0(E)$. 一个 $(E, \mathscr{E})$ 上的子 Markov 半群 (P_t) 总可以唯一地延拓为 $(E_\Delta, \mathscr{E}_\Delta)$ 上的 Markov 半群, 仍记为 (P_t).

定义 4.3.1 $(E, \mathscr{E})$ 上的转移半群 (P_t) 称为是 Feller 半群, 如果 $P_0 = I$ 且满足:

(1) 对任何 $t \geqslant 0$, $P_t(C_0(E)) \subset C_0(E)$;

(2) 对任何 $f \in C_0(E)$, $\lim_{t\downarrow 0} \|P_t f - f\| = 0$.

一个转移半群是 Feller 半群的 Markov 过程称为是 Feller 过程.

因为 $P_t 1_{E_\Delta} \equiv 1_{E_\Delta}$, 故定义中 $C_0(E)$ 可以等价地以 $C(E_\Delta)$ 代替. 另外定义条件 (2) 的一致收敛可以等价地由逐点收敛代替, 即对任何 $x \in E$, $P_t f(x) \to f(x)$. 事实上, 由 Feller 半群的性质容易验证: 对 $f \in b\mathscr{E}$, $(t, x) \mapsto P_t f(x)$ 是可测的. 我们仍然可以定义豫解算子 (U^α). 令 $L := U^\alpha(C(E_\Delta))$, 则容易验证对任何 $f \in L$, $\|P_t f - f\| \to 0$, 因此我们只需验证 L 在 $C(E_\Delta)$ 中稠, 由 Hahn-Banach 定理与 $C(E_\Delta)$ 上有界线性泛函的表示定理知这等价于 L 在 $C(E_\Delta)$ 中以逐点收敛的拓扑稠, 而 $P_t f$ 逐点收敛于 f 蕴含着当 $\alpha \to \infty$ 时, $\alpha U^\alpha f$ 逐点收敛于 f. 由豫解方程式, $C(E_\Delta)$ 在 U^α 下的像 $U^\alpha(C(E_\Delta))$ 实际上与 $\alpha > 0$ 无关, 因此对任何 $f \in C(E_\Delta)$,

存在 $f_n \in C(E_\Delta)$, 使得 $U^\alpha f_n$ 逐点收敛于 f. 另外如果 (P_t) 是 Feller 半群, 则 (i) $U^\alpha(C(E_\Delta)) \subset C(E_\Delta)$; (ii) 对任何 $f \in C(E_\Delta)$, $\lim_{\alpha\to\infty} \|\alpha U^\alpha f - f\| = 0$.

实际上, 后面我们将看到由 $\mathbf{R}^d$ 上卷积半群诱导的转移半群都是 Feller 半群. 下面我们看一个不是 Feller 半群的例子.

例 4.3.1 设 (P_t) 是 $\mathbf{R}$ 上 Brown 运动的转移半群. 令 $E := \mathbf{R}\backslash\{0\}$. 当然 E 是局部紧具有可数基度量空间. 用 Q_t 表示 P_t 限制在 E 上的核. 显然对任何 $x \in E, A \in \mathscr{B}(E)$, $Q_t(x,\ A) = P_t(x,\ A)$. 因此 (Q_t) 是 E 上的转移半群,但不是 Feller 半群. ▮

先证明下面的引理.

引理 4.3.1 设 $\mathbf{D}$ 是 $C(E_\Delta)$ 的一个可分离点的子集, h 是 $\mathbf{R}_+$ 到 E_Δ 的映射, S 是 $\mathbf{R}_+$ 的一个稠子集, 如果对任何 $g \in \mathbf{D}$, $(g\circ h)|_S$ 在 $\mathbf{R}_+$ 的每一点上有左右极限, 则 $h|_S$ 在 $\mathbf{R}_+$ 的每一点上也有左右极限.

证明 假设 $h|_S$ 在点 $x \in \mathbf{R}_+$ 没有右极限, 则存在 $\{t_n\}, \{t'_n\} \subset S$ 且 $t_n \downarrow\downarrow x$, $t'_n \downarrow\downarrow x$, 使得 $\lim h(t_n) = a \neq b = \lim h(t'_n)$. 取 $g \in \mathbf{D}$, 使得 $g(a) \neq g(b)$, 那么 $\lim g\circ h(t_n) = g(a) \neq g(b) = \lim g\circ h(t'_n)$, 与条件矛盾. □

设以 $(E, \mathscr{E})$ 为状态空间的 Markov 过程

$$X = (\Omega, \mathscr{F}, X_t, \mathbb{P}^x)$$

是 Feller 半群 (P_t) 的(§2.3的意义下)一个实现. 注意这里 $(t,\omega) \mapsto X_t(\omega)$ 不一定可测. 但是若 $f \in \mathscr{E}_+$, 则函数 $U^\alpha f$ 总是 α-上平均的且与引理 4.2.1 类似, 只要有需要的可积性, 过程 $\mathrm{e}^{-\alpha t} U^\alpha f(X_t)$ 是上鞅.

取 S 为 $\mathbf{R}_+$ 的一个可数稠密子集, 比如有理数集. 令 $g := U^\alpha f$, 其中 $f \in C(E_\Delta)$, $f \geqslant 0$, 则过程 $\mathrm{e}^{-\alpha t} g(X_t)$ 是上鞅. 前面在 §3.3 中, 我们已经证明存在 $N_g \subset \Omega$ 且对任何 $x \in E_\Delta$, $\mathbb{P}^x(N_g) = 0$, 使得对 $\omega \notin N_g$, $g(X_\cdot(\omega))|_S$ 在 $\mathbf{R}_+$ 上有左右极限. 取 $C(E_\Delta)$ 的一个分离点的非负函数组成的可列子集 L, 令 $\mathbf{D} := \{U^n(nf) : n \geqslant 1, f \in L\}$, 则由 Feller 半群的性质推出 $U^n(nf) \to f$, $f \in L$. 因此可列集 $\mathbf{D}$ 分离 E_Δ 的点, 令

$$N_1 := \bigcup_{g\in\mathbf{D}} N_g,\ \Omega_1 := N_1^c,$$

因 $\mathbf{D}$ 可列, 故 N_1 的概率等于 0, 且如果 $\omega \in \Omega_1$, 那么对任何 $g \in \mathbf{D}$, $g(X_\cdot(\omega))|_S$ 在

$\mathbf{R}_+$ 上有左右极限, 由引理, $X_\cdot(\omega)|_S$ 在 $\mathbf{R}_+$ 上有左右极限.

令 Ω_2 是满足下列条件的轨道 $\omega \in \Omega$ 全体: 对任何 $t \in S$, $X_t(\omega) \in E$ 蕴含着集合 $\{X_s(\omega): s \in S \cap [0,t]\}$ 在 E 中有界, 即包含在 E 的一个紧子集中. $N_2 := \Omega_2^c$, 则 $N_2 \subset \bigcup_{t \in S} \Gamma_t$, 其中

$$\Gamma_t = \{\omega \in \Omega: X_t(\omega) \in E \text{ 而集合 } \{X_s(\omega): s \in S \cap [0,t]\} \text{ 无界}\}.$$

取 $f \in C_0(E)$ 且 f 在 E 上严格正. 令 $g := U^1 f$, 则 $g \in C_0(E)$, $g(\Delta) = 0$ 且 g 在 E 上严格正. 因为 $g \in C_0(E)$, 故对任何 $t \in S$,

$$\Gamma_t \subset \{\mathrm{e}^{-t} g(X_t) > 0, \ \inf_{s \in S \cap [0,t]} \mathrm{e}^{-t} g(X_s) = 0\}.$$

而 $\mathrm{e}^{-t} g(X_t)|_{t \in S}$ 是非负上鞅, 由定理 3.3.4 推出 Γ_t 是零概率集.

令 $\Omega_0 := \Omega_1 \cap \Omega_2$, 则对任何 $x \in E$, $\mathbb{P}^x(\Omega_0) = 1$. 再令

$$Z_t(\omega) := \lim_{t_n \downarrow\downarrow t, t_n \in S} X_{t_n}(\omega), \ \omega \in \Omega_0.$$

显然 $Z = (Z_t)$ 是右连续且具左极限的过程, 且 Δ 是 Z 的坟墓点. 事实上, 对任何 $t \geqslant 0$, 如果 $Z_t = \Delta$, 那么对任何 $s > t$, $s \in S$, 集合 $\{X_r: r \in S \cap [0,s]\}$ 在 E 中无界, 由 Ω_2 的定义推出 $X_s = \Delta$, 因此对任何 $s > t$, $Z_s = \Delta$.

下面我们证明过程 $(\Omega, \mathscr{F}, \mathscr{F}_t, Z_t, \theta_t, \mathbb{P}^x)$ 是以 $(E, \mathscr{E}^*)$ 为状态空间的 Markov 过程, 其中 $(\mathscr{F}_t)$ 是原过程 X 的自然流经过上节所述的完备化后的 σ- 代数流.

首先 Z 是 $(\mathscr{F}_{t+})$ 适应的, 另外对任何 $t \geqslant 0$, 取 $\{t_n\} \subset S$ 并严格递减趋于 t. 任取 $s > t$ 且 $s - t \in S$, $A \in \mathscr{F}_{t+}$, $x \in E$, $f \in C(E_\Delta)$, 则 $A \in \mathscr{F}_{t_n}$, 令 $s_n := t_n - t + s$, 由 X 的 Markov 性得

$$\mathbb{E}^x(f(X_{s_n}); A) = \mathbb{E}^x(f(X_{t_n + s - t}; A) = \mathbb{E}^x[P_{s-t} f(X_{t_n}); A].$$

让 n 趋于无穷, 得到 $X_{s_n} \to Z_s$, $X_{t_n} \to Z_t$, 而 $P_{s-t} f$ 连续. 由控制收敛定理,

$$\mathbb{E}^x(f(Z_s); A) = \mathbb{E}^x[P_{s-t} f(Z_t); A].$$

现在两边关于 s 都是右连续的, 这推出上式对任何 $s \geqslant t$ 成立.

下面我们证明 Z 是 X 的一个修正, 即对任何 t, $X_t = Z_t$ a.s. 取任何 $f, g \in C(E_\Delta)$, 再取 $t_n \downarrow\downarrow t$, 由 Markov 性,

$$\mathbb{E}^x[f(X_{t_n})g(X_t)] = \mathbb{E}^x g(X_t)\mathbb{E}^{X_t} f(X_{t_n-t}) = P_t(gP_{t_n-t}f)(x),$$

当 $n \to \infty$, 右边收敛于 $P_t(gf)(x)$. (这实际上证明了 X 本身一定是随机右连续的.) 现在让 $\{t_n\} \subset S$, 则有 $\mathbb{E}^x[f(Z_t)g(X_t)] = P_t(gf)(x)$, 应用 Stone-Weierstrass 定理推出 $\mathbb{E}^x f(Z_t, X_t) = \mathbb{E}^x f(X_t, X_t)$ 对 $E_\Delta \times E_\Delta$ 上任何连续函数 f 成立, 这蕴含着 $\mathbb{P}^x(X_t = Z_t) = 1$, 即 Z 是 X 的修正, 故它们有相同的转移半群 (P_t) 且推出 $Z = (\Omega, \mathscr{F}, \mathscr{F}_t, Z_t, \mathbb{P}^x)$ 也是 Markov 过程.

这样我们证明了修正 Z 是右连续简单 Markov 过程. 因为 Feller 过程的半群和豫解把连续函数映为连续函数, 故 Feller 过程是 Borel 右过程.

例 4.3.2 Feller 半群的最自然的成员是 $\mathbf{R}^d$ 上卷积半群. 容易验证 $\mathbf{R}^d$ 上的卷积半群 $\pi = \{\pi_t\}$ 所诱导的转移半群 (P_t) 是 Feller 半群. 首先对任何连续函数 $f \in C_0(\mathbf{R}^d)$, $\pi_t * f$ 是连续的且由有界收敛定理得 $\pi * f \in C_0(\mathbf{R}^d)$, 另外对任何 $x \in \mathbf{R}^d$, 由 π 的弱连续性, $\pi_t * f$ 是点点收敛于 f. 因此存在一个右连续的 Markov 过程以 π 为卷积半群, 称为 Lévy 过程. ▮

下面给出标准过程与 Hunt 过程的定义, 它们都是很重要的 Markov 过程类.

定义 4.3.2 一个以 $(E, \mathscr{E})$ 为状态空间的 Borel 右过程

$$X = (\Omega, \mathscr{F}, \mathscr{F}_t, X_t, \theta_t, \mathbb{P}^x)$$

称为是一个标准过程, 如果 X 在 $[0, \zeta)$ 上拟左连续, 即对任何递增趋于 T 的 $(\mathscr{F}_t)$ 停时列 $\{T_n\}$, 有 $X(T_n) \longrightarrow X(T)$ a.s. 在 $\{T < \zeta\}$.

如果上面定义中的拟左连续性的收敛 a.s. 在 $\{T < \infty\}$ 上成立, 或者说 X 在 $[0, \infty)$ 上拟左连续, 则称 X 是一个 Hunt 过程(以纪念 Hunt 在概率位势理论领域的杰出贡献). 实际上, 我们可以证明标准过程的几乎所有轨道在 ζ 前具有左极限. (参考文献 [6] 的 §3.1.) 总结上述结论我们得到下面的定理.

定理 4.3.1 设 (P_t) 是局部紧具有可数基的 Hausdorff 空间 $(E, \mathscr{E})$ 上的 Feller 半群, 则存在一个标准过程 (实际上可以证明是 Hunt 过程) X, 它的转移半群恰是 (P_t).

证明 由定义, 我们只需验证上面得到的右连续具有强 Markov 性的随机过程

$$X = (\Omega, \mathscr{F}, \mathscr{F}_t, X_t, \theta_t, \mathbb{P}^x)$$

是正规的且拟左连续的. 对任何 $f \in C(E_\Delta)$, $x \in E$, 因为 X 是右连续的,

$$\mathbb{P}^x f(X_0) = \lim_{t \downarrow 0} \mathbb{P}^x(f(X_t)) = \lim_{t \downarrow 0} P_t f(x) = f(x),$$

因此 $\mathbb{P}^x(X_0 = x) = 1$.

另外取递增趋于 T 的停时列 $\{T_n\}$, 因为我们考虑在集合 $\{T < \zeta\}$, 故不妨设 T 是有界的, 对 $t > 0$, 令 $L := \lim_n X(T_n)$, $L_t := \lim_n X(T_n + t)$, 极限的存在性由轨道具有左极限的性质推出, 因为 $T_n + t \in [T, T + t]$, 故由轨道的右连续性, 不难验证当 $t \downarrow 0$ 时, $L_t \to X(T)$. 任取 $x \in E$, $f, g \in C(E_\Delta)$, 将强 Markov 性应用于 T_n,

$$\begin{aligned}
\mathbb{P}^x(f(L)g(L_t)) &= \lim_n \mathbb{P}^x[f(X(T_n))g(X(T_n + t))] \\
&= \lim_n \mathbb{P}^x[f(X(T_n))P_t g(X(T_n))] \\
&= \mathbb{P}^x[f(L)P_t g(L)],
\end{aligned}$$

有界收敛定理推出

$$\mathbb{P}^x[f(L)g(X(T))] = \mathbb{P}^x[f(L)g(L)],$$

因此我们有 $\mathbb{P}^x(L = X(T)) = 1$. □

下面我们提到 Feller 过程时, 通常是指它的标准过程的版本. 我们给出一个使标准过程的轨道连续的充分条件.

首先, 一个标准过程 X 的轨道在有限时间内是有界的, 即对任何 t, 对几乎所有的 $\omega \in \Omega$, 集合 $\{X_s(\omega) : s \in [0, t], t < \zeta(\omega)\}$ 是有界的. 事实上, 取紧集列 $\{K_n\}$ 满足 $K_{n-1} \subset K_n^\circ$, $E = \bigcup_n K_n$, 令 T_n 是 K_n^c 的首中时, 则 $\{T_n\}$ 是递增的停时列, 记其极限为 T, 由拟左连续性 $\lim_n X(T_n) = X(T)$ a.s. 在 $\{T < \zeta\}$ 上, 由轨道右连续性, $X(T_n) \notin K_n^\circ$, 那么 $X(T_n) \notin K_{n-1}$, 因此 $T = \zeta$ a.s.. 故对任何 $t < \zeta$, 存在 n, 使得 $t < T_n$, 推出当 $s \in [0, t]$ 时, $X_s \in K_n$.

定理 4.3.2　设 X 是 $(E, \mathscr{E})$ 上的标准过程, 且满足对 E 的任何紧子集 K 和 $\epsilon > 0$, 当 $t \downarrow 0$ 时, $\dfrac{1}{t} P_t(x, \{y : d(y, x) > \epsilon\})$ 对 $x \in K$ 一致地收敛于零, 则 X 的几乎所有轨道在 $[0, \zeta)$ 上连续.

证明　因为轨道在 $[0, \zeta)$ 是右连续且具有左极限的, 因此需证明对任何 $t \geqslant 0$, $\epsilon > 0$,

$x \in E$, 有

$$\mathbb{P}^x\left[\bigcap_{n=1}^{\infty}\bigcup_{k=1}^{n}\{d[X((k-1)t/n), X(kt/n)] > \epsilon\}, t < \zeta\right] = 0,$$

其中 d 是 E 上的度量, 即 X 在一个稠子集上一致连续.

取上面所言的紧集列 $\{K_n\}$ 及停时 $\{T_n\}$, 因为当 $t < T_n$ 时, 必 $X([0,t]) \in K_n$, 因此对任何 $m \geqslant 1$,

$$\begin{aligned}
&\mathbb{P}^x\left(\bigcup_{k=1}^{n}\{d[X((k-1)t/n), X(kt/n)] > \epsilon\}, t < T_m\right)\\
\leqslant\ &\mathbb{P}^x\left(\bigcup_{k=1}^{n}\{d[X((k-1)t/n), X(kt/n)] > \epsilon\}, X([0,t]) \subset K_m\right)\\
\leqslant\ &\sum_{k=1}^{n}\mathbb{P}^x(d[X((k-1)t/n), X(kt/n)] > \epsilon; X((k-1)t/n) \in K_m)\\
=\ &\sum_{k=1}^{n}\mathbb{E}^x\left(\mathbb{P}^{X((k-1)t/n)}(d(X_0, X_{t/n}) > \epsilon); X((k-1)t/n) \in K_m\right)\\
\leqslant\ &n \cdot \sup_{x \in K_m}\mathbb{P}^x(d(X_0, X_{t/n}) > \epsilon),
\end{aligned}$$

由条件它当 n趋于无穷时收敛于零. 而 $T_m \uparrow \zeta$ a.s., 故推出

$$\mathbb{P}^x\left[\bigcap_{n=1}^{\infty}\bigcup_{k=1}^{n}\{d[X((k-1)t/n), X(kt/n)] > \epsilon\}, t < \zeta\right] = 0.$$

□

例 4.3.3 设 (P_t) 是 $\mathbf{R}^d$ 上热核半群, 取任何紧子集 K 及闭子集 F, 如果 $d(K,F) > \epsilon > 0$, 则对 $x \in K$,

$$\frac{1}{t}P_t(x,F) = \int_{y\in F}\frac{1}{(2\pi t)^{\frac{d}{2}}t}\mathrm{e}^{\frac{-|y-x|^2}{2t}}\mathrm{d}y.$$

因 $|y-x| \geqslant \epsilon$, 由控制收敛定理, 上式当 $t \downarrow 0$ 时在 K 上一致趋于零. 因此, 我们实际上用另一种方法证明了 Brown 运动是一个连续标准过程.

下面我们讨论Feller 过程的生成算子. 上一节中定义的生成算子是一个等价类, 而对 Feller 过程, 我们可以在连续函数上来考虑生成算子问题. 设 X 是 Feller 过程,

(P_t) 是对应的 Feller 半群. 因为 $C_0(E)$ 是 Banach 空间, (P_t) 就是其上的一个强连续半群, 由 Hill-Yasida 定理, 它有一个在 $C_0(E)$ 上的生成算子 $(A, D(A))$,

$$Af = \lim_{t\downarrow 0} \frac{P_t f - f}{t},$$

其中 $D(A)$ 是使得上面极限是范数意义下收敛的 f 全体. 令 $(U^\alpha)_{\alpha>0}$ 是 (P_t) 的豫解算子族, 那么有下面的 Hille-Yosida 定理. 它实际上对任何 Banach 空间上的强连续压缩半群都是成立的.

定理 4.3.3 (Hille, Yosida) 设 (P_t) 是 Feller 半群, 其生成算子与豫解定义如上, 则

(1) 如果 $f \in D(A)$, 则 $P_t f \in D(A)$, $t \mapsto P_t f$ 强可微且

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t f = AP_t f = P_t Af.$$

另外

$$P_t f - f = \int_0^t P_s Af \mathrm{d}s = \int_0^t AP_s f \mathrm{d}s;$$

(2) $D(A)$ 是稠定闭算子;

(3) 对任何 $\alpha > 0$, $D(A)$ 到 $C_0(E)$ 的算子 $\alpha - A$ 是一一对应且 $U^\alpha = (\alpha - A)^{-1}$;

(4) 半群 (P_t) 由其生成算子唯一决定.

下面的定理是针对于 Feller 半群的 Markov 性的.

定理 4.3.4 Feller 半群的生成算子满足正最大值原理: 如果 $f \in D(A)$ 且 $x_0 \in E$ 满足 $0 \leqslant f(x_0) = \sup_{x\in E} f(x)$, 那么 $Af(x_0) \leqslant 0$.

证明　因为

$$P_t f(x_0) - f(x_0) = \int_E f(y) P_t(x_0, \mathrm{d}y) - f(x_0) \leqslant f(x_0)(P_t(x_0, E) - 1) \leqslant 0,$$

因此 $Af(x_0) \leqslant 0$. □

定理 4.3.5 设 $f \in D(A)$, 则过程

$$M_t^f = f(X_t) - f(X_0) - \int_0^t Af(X_s) \mathrm{d}s$$

是鞅. 特别地, 如果 $Af = 0$, 则 $f(X_t)$ 是鞅. 反之若 $f \in C_0(E)$ 且存在 $g \in C_0(E)$, 使得

$$f(X_t) - f(X_0) - \int_0^t g(X_s)\mathrm{d}s$$

是鞅, 则 $f \in D(A)$ 且 $Af = g$.

证明 先设 $f \in D(A)$. 由 Hille-Yosida 定理, 对任何 $x \in E$, $t \geqslant 0$,

$$\mathbb{E}^x(f(X_t) - f(X_0)) = P_t f(x) - f(x) = \int_0^t P_s Af(x)\mathrm{d}s = \mathbb{E}^x \int_0^t Af(X_s)\mathrm{d}s.$$

因此对任何 $t > s \geqslant 0$, 由 Markov 性有

$$\begin{aligned}\mathbb{E}^x(f(X_t) - f(X_s)|\mathscr{F}_s) &= \mathbb{E}^{X_s}(f(X_{t-s}) - f(X_0)) \\ &= \mathbb{E}^{X_s} \int_0^{t-s} Af(X_u)\mathrm{d}u \\ &= \mathbb{E}^x(\int_s^t Af(X_u)\mathrm{d}u|\mathscr{F}_s).\end{aligned}$$

反之若 $f \in C_0(E)$ 且 $f(X_t) - f(X_0) - \int_0^t g(X_s)\mathrm{d}s$ 是鞅, 则 $P_t f - f - \int_0^t P_s g\mathrm{d}s = 0$. 因此

$$\lim_{t\to 0} \frac{P_t f - f}{t} = \lim_{t\to 0} \frac{1}{t} \int_0^t P_s g\mathrm{d}s = g.$$

完成证明. □

例 4.3.4 设 (P_t) 是 $\mathbf{R}$ 上向右一致平移, 即

$$P_t(x, B) = 1_B(x + t),\ x \in \mathbf{R}, B \in \mathscr{B}(\mathbf{R}), t \geqslant 0,$$

则 $P_t f(x) = f(x + t)$. 函数 f 称为右一致可导, 如果

$$\lim_{t\downarrow 0} \frac{f(x + t) - f(x)}{t}$$

对 $x \in \mathbf{R}$ 一致存在, 那么 (P_t) 无穷小算子的定义域 $D(A)$ 是 $C_0(\mathbf{R})$ 中右一致可导函数全体, 而

$$Af(x) = D_+ f(x),\ f \in D(A), x \in \mathbf{R},$$

其中 D_+ 表示函数的右导数.

Feller 过程的最重要的组成部分就是 Lévy 过程. 它具有独立及平稳的增量, 它包括我们熟悉的大多数过程, 如 Poisson 过程, Brown 运动, 一致移动及稳定过程等. 它的美丽之处是它有一个极其简洁的分析刻画.

定义 4.3.3 以 $\mathbf{R}^d$ 为状态空间的 Feller 过程 X 称为是一个 Lévy 过程, 如果其转移半群 (P_t) 满足对任何 $t \geqslant 0, x \in \mathbf{R}^d, A \in \mathscr{B}(\mathbf{R}^d)$ 有

$$P_t(x, A) = P_t(0, A - x).$$

上面的性质等价于 X 是独立增量过程, 即等价于对任何 $t > s \geqslant 0$, $x \in \mathbf{R}^d$, $A \in \mathscr{B}(\mathbf{R}^d)$,

$$\mathbb{P}^x(X_t - X_s \in A | \mathscr{F}_s) = \mathbb{P}^x(X_{t-s} - X_0 \in A).$$

设 X 是 Lévy 过程. 对 $t \geqslant 0$, 令 $\pi_t(A) := \mathbb{P}^0(X_t \in A)$, 则概率测度族 $\pi = \{\pi_t : t \geqslant 0\}$ 对卷积具有半群性, 即对 $t, s \geqslant 0$, $\pi_t * \pi_s = \pi_{t+s}$, 因为 X_{t+s} 是独立随机变量 $X_{t+s} - X_s$ 与 X_s 的和. 令 $P_t(x, A) := \pi_t(A - x)$, $x \in \mathbf{R}^d$, $A \in \mathscr{B}(\mathbf{R}^d)$, 那么, 容易验证 (P_t) 是 $\mathbf{R}^d$ 上转移半群. 由于 Lévy 过程是右连续的, 故对任何 $f \in C_b(\mathbf{R}^d)$, 由控制收敛定理, 当 $t \downarrow 0$, $\pi_t(f) = \mathbb{E}f(X_t - X_0) \longrightarrow f(0)$.

定义 4.3.4 $\mathbf{R}^d$ 上的概率测度族 $\{\pi_t : t > 0\}$ 称为是卷积半群, 如果

(1) 对任何 $t, s > 0$, $\pi_t * \pi_s = \pi_{t+s}$;

(2) 当 $t \to 0$, π_t 弱收敛于 ϵ_0, 即对 $f \in C_b(\mathbf{R}^d)$, $\pi_t(f) \to f(0)$.

不妨补充定义 $\pi_0 = \epsilon_0$. 如果 $t \mapsto l(t)$ 是 $\mathbf{R}$ 上非线性的加群同态, 即满足可加性: $l(t+s) = l(t) + l(s)$, 则概率测度族 $\{\epsilon_{l(t)}\}$ 满足 (1) 但不满足 (2).

Lévy 过程唯一决定一个卷积半群. 反之一个 $\mathbf{R}^d$ 上对卷积具有半群性的概率测度族 $\{\pi_t : t \geqslant 0\}$ 决定一个转移半群: $P_t(x, A) := \pi_t(A - x)$, $t \geqslant 0, x \in \mathbf{R}^d, A \in \mathscr{B}(\mathbf{R}^d)$. 显然此转移半群具有空间齐性: $P_t(x, A) = P_t(0, A - x)$, 由定理 2.3.3, (P_t) 可以实现为一个时齐 Markov 过程, 由例 2.1.6 看出它是 一个平稳独立增量过程. 进一步地, 我们有下面的定理.

定理 4.3.6 Lévy 过程唯一决定一个 $\mathbf{R}^d$ 上卷积半群, 反之, 一个 $\mathbf{R}^d$ 上卷积半群可实现为一个 Lévy 过程. 实际上, 一个卷积半群诱导的转移半群是 Feller 半群.

证明 我们只需证明卷积半群诱导的半群是 Feller 半群就足够了. 设 $\{\pi_t\}$ 是 $\mathbf{R}^d$ 上

卷积半群, (P_t) 是其诱导的转移半群, 那么对任何 $f \in C_0(\mathbf{R}^d)$, $P_t f = \pi_t * f$, 它显然是有界连续的. 由有界收敛定理,

$$\lim_{x\to\infty} \pi_t f(x) = \lim_{x\to\infty} \int f(x+y)\pi_t(\mathrm{d}y) = \int \lim_{x\to\infty} f(x+y)\pi_t(\mathrm{d}y) = 0.$$

另外对任何 $x \in \mathbf{R}^d$, $\lim_{t\to 0} P_t f(x) = \lim_{t\to 0} \int f(x+y)\pi_t(\mathrm{d}y) = f(x)$. 因此 (P_t) 是 Feller 半群. □

如果 X 是 Lévy 过程, 我们总写 $\mathbb{P} := \mathbb{P}^0$. 因为它是空间齐次的, 所以 $\mathbb{P}^x = \mathbb{P}\circ\gamma_x^{-1}$, 其中 γ_x 是平移算子满足 $X_t\circ\gamma_x = X_t + x$. 可以验证 X 的强 Markov 性与空间齐性可刻画为对任何停时 T, 过程 $t \mapsto X_{t+T} - X_T = (X_t - X_0)\circ\theta_T$ 是一个与 $\mathscr{F}_T$ 独立的与 $t \mapsto X_t - X_0$ 等价的过程.

设 $\{\pi_t\}$ 是 $\mathbf{R}^d$ 上一个卷积半群, 那么 π_t 的特征函数

$$\hat{\pi}_t(\xi) := \int_{\mathbf{R}^d} \mathrm{e}^{\mathrm{i}(x,\xi)}\pi_t(\mathrm{d}x),\ \xi \in \mathbf{R}^d$$

是 $\mathbf{R}^d$ 上一个连续有界的复值函数, 且对 $t, s \geqslant 0$, $\hat{\pi}_{t+s} = \hat{\pi}_t \cdot \hat{\pi}_s$. 另外 $t \mapsto \hat{\pi}_t(\xi)$ 有右连续性, 因此存在 $\mathbf{R}^d$ 上连续复值函数 ϕ, 使得 $\hat{\pi}_t = \mathrm{e}^{-t\phi}$, $t \geqslant 0$. 函数 ϕ 称为是卷积半群的 Lévy 指数. 由特征函数的性质, 卷积半群由其 Lévy 指数唯一决定. 下面我们给出几个主要的卷积半群并计算其 Lévy 指数.

例 4.3.5 (一致移动)　设 $\alpha \in \mathbf{R}^d$, 容易验证 $\{\epsilon_{\alpha t} : t \geqslant 0\}$ 是一个卷积半群, 它的 Lévy 指数是 $-\mathrm{i}(\xi, \alpha)$. ▮

例 4.3.6 (热核半群)　热传导方程

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u$$

的解是热核

$$p(t,x) := \frac{1}{(2\pi t)^{\frac{d}{2}}}\mathrm{e}^{-\frac{|x|^2}{2t}},\ t > 0, x \in \mathbf{R}^d.$$

对 $t > 0$, 令 $b_t(A) := \int_A p(t,x)\mathrm{d}x$, $A \in \mathscr{B}(\mathbf{R}^d)$, 那么从前面的引理 2.6.1 的证明中看出 $\{b_t : t \geqslant 0\}$ 是 $\mathbf{R}^d$ 上卷积半群. 让我们来计算其 Lévy 指数, 对 $\xi \in \mathbf{R}^d$,

$$\begin{aligned}\hat{b}_t(\xi) &= \int_{\mathbf{R}^d} \frac{1}{(2\pi t)^{\frac{d}{2}}} \mathrm{e}^{-\frac{|x|^2}{2t}} \mathrm{e}^{\mathrm{i}(x,\xi)} \mathrm{d}x \\ &= \int_{\mathbf{R}^d} \frac{1}{(2\pi t)^{\frac{d}{2}}} \mathrm{e}^{-\frac{|x|^2 - 2(x,\mathrm{i}t\xi) + |\mathrm{i}t\xi|^2 - |\mathrm{i}t\xi|^2}{2t}} \mathrm{d}x \\ &= \int_{\mathbf{R}^d} \frac{1}{(2\pi t)^{\frac{d}{2}}} \mathrm{e}^{-\frac{|x-\mathrm{i}t\xi|^2}{2t} - t\frac{|\xi|^2}{2}} \mathrm{d}x = \mathrm{e}^{-t\frac{|\xi|^2}{2}},\end{aligned}$$

因此热核半群的 Lévy 指数是 $\frac{|\xi|^2}{2}$. ▮

例 4.3.7 (复合 Poisson 半群) 设 J 是 $\mathbf{R}^d$ 上的一个概率测度, $\lambda > 0$ 是一个常数, 对 $t \geqslant 0$, 令

$$p_t := \mathrm{e}^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} J^{*n},$$

其中 J^{*n} 是 n-重卷积, 而 $J^{*0} := \epsilon_0$, 容易验证 p_t 是概率, 且对 $t, s \geqslant 0$, $p_{t+s} = P_t * p_s$, 而对 $f \in C_b(\mathbf{R}^d)$, 由控制收敛定理,

$$\lim_{t \downarrow 0} p_t(f) = \lim_{t \downarrow 0} \mathrm{e}^{-\lambda t}(f(0) + \sum_{n=1}^{\infty} \frac{(\lambda t)^n}{n!} J^{*n}(f)) = f(0).$$

因此 $\{p_t\}$ 是一个卷积半群, 称为是复合 Poisson 半群, 对应的过程就是复合 Poisson 过程(参考第 97 页), λ 称为过程的强度, J 称为是跃度分布. $\{p_t\}$ 的特征函数为

$$\hat{p}_t = \mathrm{e}^{-\lambda t} \mathrm{e}^{\lambda t \hat{J}} = \mathrm{e}^{-\lambda(1-\hat{J})t},$$

其中 $\hat{J}$ 是 J 的 Fourier 变换. 因此复合 Poisson 过程的 Lévy 指数是 $\lambda(1 - \hat{J})$. 当 $d = 1$, $J = \epsilon_1$ 时, 卷积半群称为是参数为 λ 的 Poisson 半群. 这时,

$$p_t(\{n\}) = \mathrm{e}^{-\lambda t} \frac{(\lambda t)^n}{n!},\ n = 0, 1, 2 \cdots .$$ ▮

例 4.3.8 (对称稳定过程) 对 $\alpha \in (0, 2)$, 令

$$\nu(\mathrm{d}x) := \frac{\alpha 2^{\alpha-1} \Gamma(\frac{\alpha+d}{2})}{\pi^{\frac{d}{2}} \Gamma(1 - \frac{\alpha}{2})} \cdot |x|^{-\alpha-d} \mathrm{d}x,\ x \in \mathbf{R}^d,$$

则 $\phi(x) := \int(1-\cos(x,y))\nu(\mathrm{d}y) = |x|^\alpha$ 定义了一个 Lévy 指数. 对应的卷积半群称为是指标为 α 的对称稳定半群, 对应的过程称为是指标为 α 的对称稳定过程. 指标为 1 的对称稳定过程称为是对称 Cauchy 过程.

卷积半群的卷积仍然是卷积半群, 因此 Lévy 指数的和也仍然是 Lévy 指数. 下面我们讨论卷积半群的 Lévy 指数的一般表示, 即所谓的 Lévy-Khinchin 公式. 为此, 我们首先讨论一类重要的分布, 即无穷可分分布. $\mathbf{R}^d$ 上的分布 μ 称为是无穷可分的, 如果对任何 $n \geqslant 1$, 存在分布 μ_n 使得 $\mu = \mu_n^{*n}$. 对应的分布函数称为无穷可分分布函数, 而对应的特征函数称为是无穷可分特征函数. 显然一个特征函数 ϕ 无穷可分当且仅当对任何 n, 存在特征函数 ϕ_n, 使得 $\phi = \phi_n^n$. 显然, 如果 $\{\pi_t\}$ 是对卷积具有半群性, 则任何 π_t 都是无穷可分分布. 先给出一个引理, 它可以由初等的不等式 $1-\cos 2\alpha \leqslant 4(1-\cos\alpha)$ 推出. 留给读者作为练习.

引理 4.3.2　设 f 是一个特征函数, 且实部为 u, 则对 $\xi \in \mathbf{R}$,

$$1 - u(2\xi) \leqslant 4(1-u(\xi)).$$

下面的定理是重要的.

定理 4.3.7　设 $\{f_n\}$ 是特征函数列, 则函数列 f_n^n 有一个连续极限 (记为 f) 当且仅当 $n(1-f_n)$ 有连续极限(记为 ϕ). 这时候 $f = \mathrm{e}^{-\phi}$.

证明　先设 $n(1-f_n)$ 有一个连续极限 ϕ, 则 $f_n \to 1$ 且在有限区间上一致, 那么对任何 $a > 0$, 对充分大的 n, $|1-f_n| < r < 1$ 在 $[-a,a]$ 上成立. 因此由 Taylor 展开, 对 $|\xi| \leqslant a$,

$$\begin{aligned} n\log f_n(\xi) &= n\log(1-(1-f_n(\xi))) \\ &= -n(1-f_n(\xi)) - \frac{n}{2}(1-f_n(\xi))^2 - \cdots, \end{aligned}$$

右边第一项的极限为 $-\phi(\xi)$, 其余的为零. 由一致性及 a 的任意性推出 $n\log f_n$ 点点收敛于 $-\phi$, 即 $f_n^n \to \mathrm{e}^{-\phi}$.

反之设 f_n^n 有一个连续极限 f, 我们先证 f 不能有零点, 因特征函数列 $|f_n|^{2n}$ 的极限是 $|f|^2$, 故不妨设 f_n, f 是实的非负的. 因 f 连续且 f_n^n 在任何有限区间上一致收敛, 故存在 $a > 0$, f 在 $[-a,a]$ 上严格正, 那么当 n 充分大时, 在闭区间 $[-a,a]$ 上, f_n^n 正且与零有一个正距离, 因此 $-n\log f_n$, 同时 $n(1-f_n)$, 在 $[-a,a]$ 上一致有

界. 由引理 4.3.2 推出 $n(1-f_n)$ 在 $[-2a,2a]$ 上一致有界, 从而由上述 Taylor 展开推出 $-n\log f_n$ 在 $[-2a,2a]$ 上一致有界. 因此 f 在 $[-2a,2a]$ 上严格正, 重复这个过程证明了 f 在 $\mathbf{R}$ 上严格正.

现在在 $\mathbf{R}$ 的任何有限区间 I 上 f_n^n 一致收敛于连续的 f. 因 f 严格正, 故 $n\log f_n$ 在 I 上一致收敛于 $\log f$. 这蕴含着 f_n 点点收敛于 1 且在任何有限区间上一致. 再由上面的 Taylor 展开推出

$$-n\log f_n(\xi) = n(1-f_n(\xi))(1+\varDelta_n(\xi)),$$

其中 $\varDelta_n(\xi) = \dfrac{1}{2}(1-f_n(\xi)) + \dfrac{1}{3}(1-f_n(\xi))^2 + \cdots \longrightarrow 0$, 因此 $n(1-f_n(\xi)) \to -\log f$. □

这个定理有许多有用的推论. 首先如果 f 是无穷可分特征函数, 则 $f = f_n^n$, 因此 $f = \lim_n \mathrm{e}^{n(f_n-1)}$, 即 f 是复合 Poisson 特征函数的极限. 另外如果 f 是形如 f_n^n 的连续极限, 则对任何 $t \geqslant 0$, 复合 Poisson 特征函数 $\exp(tn(f_n-1))$ 有连续极限 f^t, 即 f^t 是特征函数, 推出 f 是无穷可分的且如果 f^t 对应的分布为 π_t, 那么 $\{\pi_t\}$ 是卷积半群.

推论 4.3.1 下面的结论成立:

(1) 特征函数 f 无穷可分当且仅当存在特征函数列 $\{f_n\}$ 使得 f_n^n 收敛于 f;

(2) 特征函数是无穷可分的当且仅当它是复合 Poisson 特征函数列的极限;

(3) 无穷可分特征函数列的连续极限是无穷可分的;

(4) 对任何无穷可分分布 μ, 存在唯一的卷积半群 $\{\pi_t\}$ 使得 $\pi_1 = \mu$.

于是研究 Lévy 指数的表示与研究无穷可分特征函数的表示是等价的. 下面是著名的 Lévy-Khinchin 公式.

定理 4.3.8 $\mathbf{R}^d$ 上复值函数 ϕ 是一个卷积半群的 Lévy 指数当且仅当 ϕ 可表示为

$$\phi(x) = \mathrm{i}(a,x) + \frac{1}{2}(Sx,x) + \int_{\mathbf{R}^d\setminus\{0\}} \left(1 - \mathrm{e}^{\mathrm{i}(x,y)} + \frac{\mathrm{i}(x,y)}{1+|y|^2}\right) J(\mathrm{d}y), \tag{4.3.1}$$

其中 $a \in \mathbf{R}^d$, S 是 $\mathbf{R}^d$ 上对称非负定线性算子, J 是 $\mathbf{R}^d\setminus\{0\}$ 上测度且满足

$$\int \frac{|x|^2}{1+|x|^2} J(\mathrm{d}x) < \infty,$$

称为卷积半群的 Lévy 测度. 另外 S, J 由 ϕ 唯一决定.

证明 为叙述简单, 我们不妨设 $d = 1$. 先设 ϕ 有上述表示, 则 ϕ 是连续的且对任何 $n > 1$, J 在 $\{x : |x| > \frac{1}{n}\}$ 上的限制是有限测度, 这样 ϕ 可写为

$$\phi(x) = \lim_n \left[\int_{|y|>\frac{1}{n}} (1 - \mathrm{e}^{\mathrm{i}xy}) J(\mathrm{d}y) + \mathrm{i}x \cdot \int_{|y|>\frac{1}{n}} \frac{y}{1+y^2} J(\mathrm{d}y)\right] + \mathrm{i}ax + \frac{1}{2}Sx^2.$$

而右边方括号内是一个复合 Poisson 型 Lévy 指数和一个一致平移 Lévy 指数的和, 故仍然是 Lévy 指数. 由上面的推论 (3), ϕ 也是 Lévy 指数.

下面证明 ϕ 唯一决定 S, J. 定义

$$\nu(\mathrm{d}y) := \frac{y^2}{1+y^2} J(\mathrm{d}y)$$

且 $\nu(\{0\}) := S$, 那么 ν 是一个有限测度且

$$\phi(x) = \mathrm{i}ax + \int_{\mathbf{R}} \left(1 - \mathrm{e}^{\mathrm{i}xy} + \frac{\mathrm{i}xy}{1+y^2}\right) \frac{1+y^2}{y^2} \nu(\mathrm{d}y),$$

其中上面的积分函数在 $y = 0$ 处是可去连续的, 定义此处为其极限值 $\frac{x^2}{2}$, 使函数连续. 直观地如果 ϕ 二阶可微, 则二阶导数是 $(1 + x^2)\nu(\mathrm{d}x)$ 的 Fourier 变换, 唯一地决定 ν. 一般地我们定义 ϕ 的二阶差分

$$\begin{aligned}
\theta(x) &:= \int_0^1 \left[\phi(x) - \frac{1}{2}(\phi(x+h) + \phi(x-h))\right] \mathrm{d}h \\
&= -\int_0^1 \int_{\mathbf{R}} \mathrm{e}^{\mathrm{i}xy}(1 - \cos hy) \frac{1+y^2}{y^2} \nu(\mathrm{d}y)\mathrm{d}h \\
&= -\int_{\mathbf{R}} \mathrm{e}^{\mathrm{i}xy} \frac{1+y^2}{y^2} \nu(\mathrm{d}y) \int_0^1 (1 - \cos hy)\mathrm{d}h \\
&= -\int_{\mathbf{R}} \mathrm{e}^{\mathrm{i}xy} \frac{1+y^2}{y^2} (1 - \frac{\sin y}{y}) \nu(\mathrm{d}y),
\end{aligned}$$

令 $k(y) := \frac{1+y^2}{y^2}(1 - \frac{\sin y}{y})$, 则存在 $c > 0$, 使得 $c \leqslant k(y) \leqslant \frac{1}{c}$. 因此 $k \cdot \nu$ 仍然是有

限测度且其 Fourier 变换是 θ, 推出它由 θ 唯一决定, 故 ν 由 θ 从而由 ϕ 唯一决定.

现在设 ϕ 是 Lévy 指数, 则 ϕ 是连续的且由定理 4.3.7, 存在特征函数列 $\{f_n\}$, 对应分布列 $\{\mu_n\}$, 使得 $n(1-f_n) \to \phi$, 则

$$\begin{aligned}\phi(x) &= \lim_n n \int_{\mathbf{R}} (1-\mathrm{e}^{\mathrm{i}xy})\mu_n(\mathrm{d}y) \\ &= \lim_n \left[n\int_{\mathbf{R}} \left(1-\mathrm{e}^{\mathrm{i}xy}+\frac{\mathrm{i}xy}{1+y^2}\right)\mu_n(\mathrm{d}y) - \mathrm{i}x\cdot n\int_{\mathbf{R}}\frac{y}{1+y^2}\mu_n(\mathrm{d}y)\right].\end{aligned}$$

令

$$a_n := -n\int_{\mathbf{R}}\frac{y}{1+y^2}\mu_n(\mathrm{d}y),\ \nu_n(\mathrm{d}y) := \frac{ny^2}{1+y^2}\mu_n(\mathrm{d}y),\ \phi_n := n(1-f_n).$$

我们再用二阶差分,

$$\theta_n(x) := \int_0^1 \left[\phi_n(x) - \frac{1}{2}(\phi_n(x+h)+\phi_n(x-h))\right]\mathrm{d}h = -\int_{\mathbf{R}} \mathrm{e}^{\mathrm{i}xy}k(y)\nu_n(\mathrm{d}y).$$

因 ϕ_n 有连续极限, 故 θ_n 也有连续极限, 记为 θ. 由 Fourier 变换的连续性, $k\cdot\nu_n$ 弱收敛, 推出 ν_n 弱收敛于一个有限测度记为 ν, 那么

$$\int_{\mathbf{R}}\left(1-\mathrm{e}^{\mathrm{i}xy}+\frac{\mathrm{i}xy}{1+y^2}\right)\frac{1+y^2}{y^2}\nu_n(\mathrm{d}y) \longrightarrow \int_{\mathbf{R}}\left(1-\mathrm{e}^{\mathrm{i}xy}+\frac{\mathrm{i}xy}{1+y^2}\right)\frac{1+y^2}{y^2}\nu(\mathrm{d}y).$$

同时推出 a_n 收敛于某实数 a. □

定理中的 Lévy 测度所满足的条件等价于对某个(因此对所有) $\epsilon > 0$, 有

$$\int_{|x|<1}|x|^2 J(\mathrm{d}x) < \infty,\ J(\{x : |x| > \epsilon\}) < \infty.$$

因此 Lévy 指数的表达式的形式通常也写成为

$$\phi(x) = \mathrm{i}(a,x) + \frac{1}{2}(Sx,x) + \int_{\mathbf{R}^d\setminus\{0\}} \left(1-\mathrm{e}^{\mathrm{i}(x,y)}+\mathrm{i}(x,y)1_{\{|y|<1\}}\right)J(\mathrm{d}y), \qquad (4.3.2)$$

因为不难验证

$$\int \left|\mathrm{i}(x,y)\left[\frac{1}{1+|y|^2} - 1_{\{|y|<1\}}\right]\right| J(\mathrm{d}y) < \infty.$$

在这两个表达式中 a 的取值可能是不同的. 当然还可以有许多其他的等价表达式. 定理也说明 Lévy 过程的所有信息都嵌入在其 Lévy 指数上. 对这方面感兴趣的读者可阅读文献 [19], [4] 和 [2].

定理 4.3.8 说明形式如 (4.3.1) 或 (4.3.2) 的函数一定是某个 Lévy 过程的 Lévy 指数. 下面我们给出 Lévy 指数所对应过程的 Itô 分解. 让我们设 ϕ 如表达式 (4.3.2), $B=(B_t)$ 是 $\mathbf{R}^d$ 上标准 Brown 运动. 令 $X_t^{(1)} := \sqrt{S}B_t - at$, 容易计算

$$\mathbb{E}\,\mathrm{e}^{\mathrm{i}(x,X_t^{(1)})} = \mathrm{e}^{-t[\mathrm{i}(a,x)+\frac{1}{2}(Sx,x)]}.$$

另外 $J_1 := J|_{\{|x|\geqslant 1\}}$ 是有限测度, 由例 4.3.7, 存在与 B 独立的复合 Poisson 过程 $X^{(2)}$, 使得

$$\mathbb{E}\,\mathrm{e}^{\mathrm{i}(x,X_t^{(2)})} = \mathrm{e}^{-t\int(1-\mathrm{e}^{\mathrm{i}(x,y)})J_1(\mathrm{d}y)}.$$

最后 $J_2 := J|_{\{|x|<1\}}$ 是 σ- 有限测度, 由定理 2.5.4, 存在独立于 $X^{(1)}$, $X^{(2)}$ 的以 J_2 为特征测度的 $\mathbf{R}^d$ 上的平稳 Poisson 点过程 p. 对任何 $0<r<1$, 令

$$X_t^{(3,r)} := \sum_{s\leqslant t} 1_{\{r<|p(s)|<1\}}p(s) - t\int_{\mathbf{R}^d} x1_{\{r<|x|<1\}}J(\mathrm{d}x),\ t\geqslant 0.$$

右边有意义因为

$$\mathbb{E}\sum_{s\leqslant t} 1_{\{r<|p(s)|<1\}}|p(s)| = t\int_{\{x:r<|x|<1\}}|x|J(\mathrm{d}x) < \infty.$$

那么 $X^{(3,r)}$ 是右连续的且由 Poisson 点过程的性质推出它是一个鞅. 同样可以验证

$$\mathbb{E}\,\mathrm{e}^{\mathrm{i}(x,X_t^{(3,r)})} = \mathrm{e}^{-t\int(1-\mathrm{e}^{\mathrm{i}(x,y)}+\mathrm{i}(x,y))1_{\{r<|y|<1\}}J(\mathrm{d}y)}.$$

利用 Doob 的鞅不等式, 定理 3.3.5(2), 推出对任何 $t>0$ 和 $r'<r$,

$$\mathbb{E}(\sup_{s\leqslant t}|X_s^{(3,r)} - X_s^{(3,r')}|^2) \leqslant 4t\int_{\mathbf{R}^d}|x|^2 1_{\{r'<|x|<r\}}J(\mathrm{d}x).$$

因此当 $r\to 0$, 右边的极限为零. 因此 $X^{(3,r)}$ (至少一个子列)几乎处处地在任何有限区间 $[0,t]$ 上一致收敛于一个过程, 记为 $X^{(3)}$. 由此推出 $X^{(3)}$ 也是一个 Lévy 过程, 且

$$\mathbb{E}\,\mathrm{e}^{\mathrm{i}(x,X_t^{(3)})} = \mathrm{e}^{-t\int(1-\mathrm{e}^{\mathrm{i}(x,y)}+\mathrm{i}(x,y))J_2(\mathrm{d}y)}.$$

令 $X := X^{(1)} + X^{(2)} + X^{(3)}$, 则 X 是一个 Lévy 指数为 ϕ 的 Lévy 过程, 此分解称为是 Itô 分解. 粗略地说, Lévy 过程有三部分组成: 连续部分, 大跳部分与小跳部分. 但是实际上, 小跳部分中包含有漂移. 因此一般地, 连续部分与小跳部分是无法严格区分的.

如果 X 是一个 Lévy 测度为 J 的 Lévy 过程, 那么由例 2.5.1 得知 X 的跳跃过程 $p_t = \Delta X_t$ 是一个平稳 Poisson 点过程. 因为 X 与上面有 Itô 分解的 Lévy 过程是等价的, 故它们的各自的跳跃点过程有相同的特征测度, 也就是说 p 的特征测度就是 Lévy 测度. 下面我们给出 Poisson 过程的一个刻画.

定理 4.3.9 一个 Lévy 过程是 Poisson 过程当且仅当它的几乎每条轨道仅以跳跃度为 1 增加.

证明 先设 N 是 Poisson 过程. 对任何 $T > 0$,

$$\sup_{0\leqslant t\leqslant T}(N_t - N_{t-}) = \lim_n \max_{1\leqslant k\leqslant n}(N_{\frac{kT}{n}} - N_{\frac{(k-1)T}{n}}) \text{ a.s.},$$

而

$$\mathbb{P}(\max_{1\leqslant k\leqslant n}(N_{\frac{kT}{n}} - N_{\frac{(k-1)T}{n}}) \leqslant 1) = [\mathbb{P}(N_{\frac{T}{n}} \leqslant 1)]^n = \mathrm{e}^{-\lambda T}\left(1+\lambda\frac{T}{n}\right)^n,$$

当 $n \to \infty$ 时, 其极限为 1. 因此 $\mathbb{P}(\sup_{0\leqslant t\leqslant T}(N_t - N_{t-}) \leqslant 1) = 1$.

反之, 如果 N 是一个仅以跳跃度为 1 增加的 Lévy 过程, 则用 S_n 表示 N 的第 n 次跳跃的时间, $n \geqslant 1$, 那么 S_n 是停时且 $S_{n+1} = S_n + S\circ\theta_{S_n}$, 其中 $S = S_1$. 由强 Markov 性和空间齐性, 对任何 $t > 0$,

$$\mathbb{P}(S_{n+1} - S_n > t|\mathscr{F}_{S_n}) = \mathbb{P}(N_{S_n+t} - N_{S_n} = 0|\mathscr{F}_{S_n}) = \mathbb{P}(N_t = 0) = \mathbb{P}(S > t).$$

由定理 2.5.3, 我们需证明 S 是指数分布. 对任何 $t, s > 0$, 由 Markov 性, 因为在 $t < S$ 时, $N_t = 0$, 故

$$\mathbb{P}(S > t+s) = \mathbb{P}(S > t, S\circ\theta_t > s) = \mathbb{E}(\mathbb{P}^{N_t}(S > s), S > t) = \mathbb{P}(S > s)\mathbb{P}(S > t).$$

即 S 是指数分布的. □

习 题

1. 设 B 是 $\mathbf{R}$ 上 Brown 运动. 证明:

(1) $|B|$ 是 Feller 过程;

(2) $(B_{t\wedge T})$ 是 Feller 过程, 其中 T 是 $\{0\}$ 的首中时.

2. 一个转移半群 (P_t) 总是有界可测函数空间上的压缩半群, 其中在依范数意义下 $\lim_{t\to 0} P_t f = f$ 的 f 全体称为 (P_t) 的强连续中心, 记为 B_0. 设 (P_t) 是度量空间 (E,d) 上的一个转移半群. 令对 $\epsilon > 0$, 定义

$$\alpha_\epsilon(t) := \sup_{x\in E, s\leqslant t} P_s(x, \{y : d(y,x) \geqslant \epsilon\}).$$

称半群是一致连续的, 如果对任何 $\epsilon > 0$, $\lim_{t\downarrow 0} \alpha_\epsilon(t) = 0$. 证明: 如果半群是一致连续的, 那么 $C_u(E) \subset \mathbf{B}_0$, 其中 $C_u(E)$ 是 E 上有界一致连续函数全体, $\mathbf{B}_0$ 是 (P_t) 的强连续中心. 反之, 如果 $C_b(E) \subset \mathbf{B}_0$, 那么半群是一致连续的.

3. 半群 (P_t) 一致连续当且仅当对应的 Markov 过程是一致随机连续的, 即对任何 $\epsilon > 0, \epsilon' > 0$, 存在 $\delta > 0$, 使得 $P^x(d(X_t, X_s) \geqslant \epsilon) < \epsilon'$ 对所有 $x \in E$, $|t-s| < \delta$, $t, s \geqslant 0$ 成立.

4. 证明: $\mathbf{R}$ 上右一致连续的函数是一致连续的.

5. 设 $C_0^{(2)}$ 是 $C_0(\mathbf{R}^d)$ 上二次可微且其一阶与二阶导数都在 C_0 中的实值函数全体, A 是热核半群的无穷小算子. 证明: $C_0^{(2)} \subset D_A$ 且 $Af = \frac{1}{2}\triangle f$.

6. 证明: 当 $d=1$ 时, $D(A) = C_0^{(2)}$.

7. 求 Poisson 半群与复合 Poisson 半群的无穷小算子.

8. 证明: 0 点反射的 Brown 运动与 $[0,1]$ 上反射 Brown 运动都是 Feller 过程.

9. 证明: $\mathbf{R}^d$ 上的 Feller 过程 X 是 Lévy 过程当且仅当对任何 $t > s \geqslant 0$, $x \in \mathbf{R}^d$, $A \in \mathscr{B}(\mathbf{R}^d)$,

$$\mathbb{P}^x(X_t - X_s \in A | \mathscr{F}_s) = \mathbb{P}^x(X_{t-s} - X_0 \in A).$$

10. 设 X 是 Lévy 过程, 证明: 对任何 $t > 0$, $\mathbb{P}(X_t = X_{t-}) = 1$.

11. 设 $d=1$, ϕ 是 Lévy 指数, 证明:

$$\lim_{|x|\to\infty} \frac{\phi(x)}{x^2} = \frac{S}{2}.$$

12. Lévy 指数 ϕ 有界当且仅当 ϕ 是复合 Poisson 半群的 Lévy 指数.

13. 设 X 是 Lévy 指数如 (4.3.2) 的 Lévy 过程. 证明: 对任何 $a > 0$, 测度 $\dfrac{1}{t}\mathbb{P}(X_t \in \mathrm{d}x)$ 当 $t \downarrow 0$ 时在 $\{|x| > a\}$ 上淡收敛于 $J(\mathrm{d}x)$.

14. 证明: 一个平稳独立增量过程 X 对应的对卷积有半群性的测度族 $\{\pi_t : t \geqslant 0\}$ 是一个卷积半群当且仅当 X 在 0 点随机连续.

15. 设 (P_t) 是 $\mathbf{R}^d$ 上由一个卷积半群诱导的转移半群, 记 $C_0 := C_0(\mathbf{R}^d)$ 是无穷远处趋于零的连续函数全体. 证明:

(1) 对任何 $t > 0$, $P_t(C_0) \subset C_0$;

(2) 对任何 $f \in C_0$, 当 $t \downarrow 0$ 时, $P_t f$ 一致收敛于 f.

*16. 以 $\mathbf{R}$ 为状态空间且轨道是递增函数的 Lévy 过程称为是从属子(subordinator), 从属子的卷积半群是支撑在 $[0, \infty)$ 上的. 证明: $\{\eta_t\}$ 是从属子的卷积半群当且仅当存在 $b \geqslant 0$ 与一个 $(0, \infty)$ 上的测度 μ, 满足 $\int_0^\infty \dfrac{s}{1+s}\mu(\mathrm{d}s) < \infty$, 使得对 $s \geqslant 0$,

$$\int_0^\infty \mathrm{e}^{-sx}\eta_t(\mathrm{d}s) = \exp\left(-t(bx + \int_0^\infty (1 - \mathrm{e}^{-sx})\mu(\mathrm{d}x))\right).$$

17. 设 (P_t) 是 $(E, \mathscr{E})$ 转移半群且对 $A \in \mathscr{E}$, $(t, x) \mapsto P_t(x, A)$ 可测, (η_t) 是一个从属子的卷积半群. 证明:

$$P_t'(x, A) := \int_0^\infty P_u(x, A)\eta_t(\mathrm{d}u)$$

定义了 $(E, \mathscr{E})$ 上的一个转移半群且若 (P_t) 是时间齐次的, 则 (P_t') 也是.

18. 任取 $\beta \in (0, 1)$. 证明:

(1) 对任何 $t > 0$, 函数 $x \mapsto \mathrm{e}^{-tx^\beta}$, $x \in (0, \infty)$ 是完全单调的. 因此由 Bernstein 定理, 它是 $(0, \infty)$ 上唯一的概率测度 η_t^β 的 Laplace 变换.

(2) $(\eta_t^\beta : t > 0)$ 是 $(0, \infty)$ 上卷积半群, 对应的 Lévy 过程称为是 β- 单边稳定过程.

(3) 取 $\mathbf{R}^d$ 上热核半群 (b_t), 定义

$$\pi_t^\beta := \int b_u \eta_t^\beta(\mathrm{d}u),$$

则 (π_t^β) 是指标为 2β 的对称稳定半群.

19. 如果 X 是 Lévy 过程, 证明: 对任何实数 u, $\exp(\mathrm{i}uX_t + t\phi(u))$ 是复值鞅.

* 20. 设 $\{\pi_t\}$ 是卷积半群, G_0 是 $\bigcup_t \operatorname{supp}\pi_t$ 生成的最小闭子群. 证明: $G_0^\perp = \{x : \phi(x) = 0\}$, 其中 $G_0^\perp := \{y \in \mathbf{R}^d : \ \mathrm{e}^{i(x,y)} = 1, \ x \in G_0\}$.

21. 设 $\{\pi_t\}$ 是卷积半群, 其 Lévy 指数为 ϕ. 非负 Radon 测度 μ 称为是 $\{\pi_t\}$ 的不变测度, 如果对任何 t, $\mu * \pi_t = \mu$. 证明:

(1) 概率测度不可能是不变测度. Lebesgue 测度及其常数倍总是不变测度. 如果卷积半群没有其他不变测度, 说它有唯一不变测度.

(2) 对称卷积半群有唯一不变测度当且仅当 ϕ 有唯一零点. 因此 Brown 运动与对称稳定过程有唯一不变测度.

§4.4 Brown 运动与经典位势

在 §2.6 中, 我们已经证明了 Brown 运动作为以热核作为转移函数的连续轨道的 Markov 过程的存在性, Brown 运动是当然的最重要的随机过程. Brown 运动是一个 Lévy 过程, 因此是连续的强 Markov 过程. 下面我们将利用强 Markov 性研究 Brown 运动与位势理论间的关系.

设 $(\mathscr{F}_t)$ 是 Brown 运动的被强化了的自然流. T 是 $(\mathscr{F}_{t+})$-停时, 则对任何 $(\Omega, \mathscr{F})$ 有界或非负随机变量 Y,

$$\mathbb{E}^x(Y \circ \theta_T; T < \infty) = \mathbb{E}^x(\mathbb{E}^{B_T}(Y); T < \infty). \tag{SM}$$

强 Markov 性有一个更一般的形式: 设 φ 是 $[0, \infty) \times \Omega$ 上的可测有界或非负函数, T 是停时, 那么

$$\mathbb{E}^x\left(\varphi(T, \theta_T(\cdot)); T < \infty\right) = \mathbb{E}^x\left([\mathbb{E}^{B_T}\varphi(t, \cdot)]_{t=T}; T < \infty\right). \tag{SM$'$}$$

从 (SM) 推出 (SM$'$) 的证明是标准的, 读者可作为习题证明之.

强 Markov 性是非常直观且有用的性质, 下面我们给出它的一些应用. 首先我们将证明重要的反射原理. 下面我们总记 $\mathbb{P} = \mathbb{P}^0$, T_a 为点 $\{a\} \subset \mathbf{R}$ 的首中时.

定理 4.4.1 (Reflection principle) 设 $a > 0$, $t > 0$, 则

$$\mathbb{P}(T_a < t) = 2\mathbb{P}(B_t > a).$$

证明 令 $\varphi(s,\omega) := 1_{\{s<t, B_{t-s}(\omega)>a\}}$, 那么

$$\varphi(T_a, \theta_{T_a}(\cdot)) = 1_{\{T_a<t, B_t>a\}}.$$

利用强 Markov 性(SM′),

$$\begin{aligned}
\mathbb{P}(B_t > a) &= \mathbb{P}(T_a < t, B_t > a) \\
&= \mathbb{E}(\mathbb{E}^{B_{T_a}}(\varphi(s,\cdot))|_{s=T_a}, T_a < t) \\
&= \mathbb{E}(\mathbb{P}^a(B_{t-s} > a)|_{s=T_a}, T_a < t) \\
&= \frac{1}{2}\mathbb{P}(T_a < t),
\end{aligned}$$

第三个等号是因为由 Brown 运动的对称性推出 $\mathbb{P}^a(B_{t-s} > a) = \frac{1}{2}$. 证毕. □

直观地讲, 反射原理是说 $\mathbb{P}(T_a < t, B_t > a) = \mathbb{P}(T_a < t, B_t < a)$, 即 从 0 出发在 t 时刻前通过 a 的轨道中, 由反射可看出在 t 时刻在 a 上的轨道与在 t 时刻在 a 下的轨道是一一对应的, 这来自于对离散时对称随机游动轨道的观察, 离散时的证明是简单而直接的. (参考文献 [14].) 连续的情况要用到强 Markov 性, 由此也可以感觉到离散和连续时思想是类似的. 利用反射原理我们来计算 0 点出发首次通过时 T_a 的分布. 继续上面定理的计算,

$$\mathbb{P}(T_a < t) = 2\mathbb{P}(B_t > a) = 2\int_a^\infty \frac{1}{\sqrt{2\pi t}}\mathrm{e}^{-\frac{x^2}{2t}}\mathrm{d}x = \int_0^t \frac{a}{\sqrt{2\pi x^3}}\mathrm{e}^{-\frac{a^2}{2x}}\mathrm{d}x,$$

那么 T_a 的 Laplace 变换为

$$\begin{aligned}
\mathbb{E}\mathrm{e}^{-sT_a} &= 2s\int_0^\infty \mathrm{e}^{-st}\mathrm{d}t\int_a^\infty \frac{1}{\sqrt{2\pi t}}\mathrm{e}^{-\frac{x^2}{2t}}\mathrm{d}x \\
&= 2s\int_a^\infty \mathrm{d}x\int_0^\infty \mathrm{e}^{-st}\frac{1}{\sqrt{2\pi t}}\mathrm{e}^{-\frac{x^2}{2t}}\mathrm{d}t \\
&= 2s\int_a^\infty \frac{\mathrm{e}^{-\sqrt{2s}x}}{\sqrt{2s}}\mathrm{d}x = \mathrm{e}^{-\sqrt{2s}a},
\end{aligned}$$

第三个等式的关键是恒等式

$$\int_0^\infty \mathrm{e}^{-a(t^2+\frac{1}{t^2})}\mathrm{d}t = \sqrt{\frac{\pi}{4a}}\mathrm{e}^{-2a}.$$

由此可以推出过程 $a \mapsto T_a$ 是一个递增的平稳独立增量过程.

回忆 x 是 $A \in \mathscr{E}^e$ 的正则点, 如果 $\mathbb{P}^x(T_A = 0) = 1$. 下面定理证明一维 Brown 运动的点是其本身的正则点且其轨道的零点集是一个拓扑 Cantor 集.

定理 4.4.2 设 B 是 1-维 Brown 运动.

(1) 对任何 $a \in \mathbf{R}$, $\mathbb{P}^a(T_a = 0) = 1$;

(2) 对几乎所有的 ω, 轨道的零点集 $\{t : B_t(\omega) = 0\}$ 是一个 Lebesgue 测度为零的拓扑 Cantor 集, 精确地说, 它是无处稠的且无孤立点的闭集.

证明 (1) 证明的方法很多. 由 T_a 的 Laplace 变换推出对任何 $x \neq a$,

$$\mathbb{E}^x\mathrm{e}^{-sT_a} = \mathrm{e}^{-\sqrt{2s}|x-a|}.$$

令 $T = T_0$. 对 $t > 0$, 用 Markov 性, 因为 $B_t \neq 0$ a.s.,

$$\mathbb{E}\mathrm{e}^{-s(t+T\circ\theta_t)} = \mathrm{e}^{-st}\mathbb{E}(\mathbb{E}^{B_t}\mathrm{e}^{-sT}) = \mathrm{e}^{-st}\mathbb{E}\mathrm{e}^{-\sqrt{2s}|B_t|}.$$

让 $t \downarrow 0$, 得 $\mathbb{E}\mathrm{e}^{-sT} = 1$, 因此 $\mathbb{P}(T = 0) = 1$.

一个更简单的方法是利用 Brown 运动的对称性和连续性. 对任何 $t > 0$, $\mathbb{P}(T_{[0,\infty)} \leqslant t) \geqslant \mathbb{P}(B_t \geqslant 0) = \dfrac{1}{2}$. 由 Blumenthal 0-1 律, $\mathbb{P}(T_{[0,\infty)} = 0) = 1$. 同理 $\mathbb{P}(T_{(-\infty,0]} = 0) = 1$. 由连续性推出 $\mathbb{P}(T = 0) = 1$.

(2) 零点集是闭且 Lebesgue 测度为零是显然的事实, 因此它也一定是无处稠的. 现在证明没有孤立的零点. 固定 $0 < s < t$, 令 $T_s := s + T\circ\theta_s$, 如果 B 在 (s,t) 间恰有一个零点, 那么 $t > T_s$ 表示 (s,t) 间至少有一个零点, 且 $T\circ\theta_{T_s} > 0$ 表示 T_s 后不能再马上命中 0, 因此若用 $N_{s,t}$ 表示事件: B 在 (s,t) 仅有一个零点, 则

$$\mathbb{P}(N_{s,t}) \leqslant \mathbb{P}(T_s < t, T\circ\theta_{T_s} > 0) = \mathbb{E}(\mathbb{P}^{B_{T_s}}(T > 0); T_s < t) \leqslant \mathbb{P}(T > 0) = 0.$$

令 $N := \bigcup_{s<t,s,t\in Q} N_{s,t}$, 其中 Q 是有理数集, 那么 $\mathbb{P}(N) = 0$, 并且若 $\omega \notin N$, $t \mapsto B_t(\omega)$ 的零点没有孤立点. 证毕. □

下面我们来讨论 Brown 运动的极集与常返集. 极集的定义见定义 4.2.4. 它表

示一个过程永远不会到达的集合, 但常返集恰好相反.

定义 4.4.1 Borel 集 A 称为是常返的, 如果对任何 x 有 $\mathbb{P}^x(T_A < \infty) = 1$. 否则称为暂留的.

也就是说极集是几乎所有轨道都不会遇到的集合, 而常返是几乎所有轨道一定会遇到的集合. 常返的概念类似于 Markov 链中所定义的. 首先上面 T_a 的 Laplace 变换的表达式说明 $\mathbb{P}(T_a < \infty) = 1$, 即证明了一维 Brown 运动的单点集因此任何非空集都是常返的. 但是二维及以上的 Brown 运动的单点集却是极集. 这主要是因为

$$\int_0^1 p(t,0)\mathrm{d}t = \int_0^1 \frac{1}{(2\pi t)^{\frac{d}{2}}}\mathrm{d}t = \begin{cases} < \infty, & d = 1, \\ = \infty, & d \geqslant 2. \end{cases}$$

另外容易验证

$$\int_0^\infty p(t,x)\mathrm{d}t = \int_0^\infty \frac{1}{(2\pi t)^{\frac{d}{2}}}\mathrm{e}^{-\frac{|x|^2}{2t}}\mathrm{d}t = \begin{cases} +\infty, & d = 1,2, \\ \dfrac{C(d)}{|x|^{d-2}}, & d \geqslant 3, \end{cases}$$

其中 $C(d)$ 是只与 d 有关的常数. 这决定了在 $d \leqslant 2$ 时紧集是常返的而 $d \geqslant 3$ 时是暂留的.

定理 4.4.3 设 B 是 d-维 Brown 运动.

(1) 当 $d \geqslant 2$ 时, 单点集是极集.

(2) 当 $d = 2$ 时, 不是极集的 Borel 集一定是常返的. 因此任何开集是常返的.

(3) 当 $d \geqslant 3$ 时, 紧集是暂留的.

证明 (1) 设 T_0 是 0 的首中时, 对 $M > s > 0$, 令 $T^{(s)} := s + T_0 \circ \theta_s$, 那么 $T^{(s)}$ 是停时. 用强 Markov 性, 对任何 $x \in \mathbf{R}^d$ 和以原点为中心 r 为半径的球 G_r,

$$\begin{aligned}
&\mathbb{E}^x\left[\left(\int_0^1 1_{G_r}(B_t)\mathrm{d}t\right)\circ\theta_{T^{(s)}}, T^{(s)} \leqslant M\right] \\
&= \mathbb{E}^x\left(\mathbb{E}^{B(T^{(s)})}\left(\int_0^1 1_{G_r}(B_t)\mathrm{d}t\right), T^{(s)} \leqslant M\right) \\
&= \mathbb{P}^x(T^{(s)} \leqslant M)\int_{G_r}\mathrm{d}y\int_0^1 p(t,y)\mathrm{d}t.
\end{aligned}$$

另一方面,

$$\begin{aligned}
&\mathbb{E}^x\left[\left(\int_0^1 1_{G_r}(B_t)\mathrm{d}t\right)\circ\theta_{T^{(s)}}, T^{(s)}\leqslant M\right]\\
=&\ \mathbb{E}^x\left(\int_0^1 1_{G_r}(B_{t+T^{(s)}})\mathrm{d}t, T^{(s)}\leqslant M\right)\\
=&\ \mathbb{E}^x\left(\int_{T^{(s)}}^{1+T^{(s)}} 1_{G_r}(B_t)\mathrm{d}t, T^{(s)}\leqslant M\right)\\
\leqslant&\ \mathbb{E}^x\left(\int_s^{M+1} 1_{G_r}(B_t)\mathrm{d}t\right)\\
=&\int_{G_r}\mathrm{d}y\int_s^{M+1}p(t,y-x)\mathrm{d}t\leqslant |G_r|\cdot\frac{M+1-s}{(2\pi s)^{\frac{d}{2}}},
\end{aligned}$$

其中 $|G_r|$ 表示 G_r 的体积. 因此有

$$\mathbb{P}^x(T^{(s)}\leqslant M)\int_{G_r}\mathrm{d}y\int_0^1 p(t,y)\mathrm{d}t\leqslant |G_r|\cdot\frac{M+1-s}{(2\pi s)^{\frac{d}{2}}}.$$

由 Fatou 引理, $\lim_{y\to 0}\int_0^1 p(t,y)\mathrm{d}t=\infty$, 故推出对任何 $M>s>0$, $\mathbb{P}^x(T^{(s)}\leqslant M)=0$. 让 $M\to\infty$, $s\to 0$, 便推出 $\mathbb{P}^x(T<\infty)=0$.

(2) 设 $f(x):=\mathbb{P}^x(T_A<\infty)$, 则 $P_tf=\mathbb{P}^{\cdot}(T_A\circ\theta_t<\infty)\leqslant f$ 且 $P_tf\uparrow f$. 对任何 $n\geqslant 1$, $t>0$,

$$\begin{aligned}
\int_0^n P_s(f-P_tf)\mathrm{d}s&=\int_0^n P_sf\mathrm{d}s-\int_t^{n+t}P_sf\mathrm{d}s\\
&=\int_0^t P_sf\mathrm{d}s-\int_n^{n+t}P_sf\mathrm{d}s\leqslant t.
\end{aligned}$$

因为当 $d\leqslant 2$ 时有 $\int_0^\infty p(t,x)\mathrm{d}t=+\infty$, 故 $f=P_tf$ a.e.. 因此对 $t,s>0$, $P_tf=P_sf$ a.e. 因两边都是连续的, 故有 $P_tf=P_sf=f$. 另外由控制收敛定理,

$$\lim_{t\to\infty}|P_tf(x)-P_tf(y)|\leqslant\lim_{t\to\infty}\int|p(t,y-z)-p(t,x-z)|dz=0.$$

因此 f 是常数.

现在对 $t>0$, 由 Markov 性,

$$\begin{aligned}\mathbb{P}^y(t<T_A<\infty)&=\mathbb{P}^y(T_A\circ\theta_t<\infty,t<T_A)\\&=\mathbb{E}^y(f(B_t),t<T_A)=\mathbb{P}^y(t<T_A)\cdot f.\end{aligned}$$

让 $t\to\infty$ 得 $0=(1-f)f$, 因此 f 恒为 1 或恒为零, 即 A 是常返或是极集.

(3) 当 $d\geqslant 3$ 时, $g:=\int_0^\infty p(t,\cdot)\mathrm{d}t$ 是一个严格正局部可积函数. 取单位球 G 及紧集 A, 则由强 Markov 性,

$$\begin{aligned}\int_{G-x}g(y)\mathrm{d}y&=\mathbb{E}^x\int_0^\infty 1_G(B_t)\mathrm{d}t\\&\geqslant\mathbb{E}^x\Big(\int_{T_A}^\infty 1_G(B_t)\mathrm{d}t,T_A<\infty\Big)\\&=\mathbb{E}^x\Big(\mathbb{E}^{B(T_A)}\int_0^\infty 1_G(B_t)\mathrm{d}t,T_A<\infty\Big)\\&\geqslant\mathbb{P}^x(T_A<\infty)\inf_{z\in\overline{A}}\int_{G-z}g(y)\mathrm{d}y.\end{aligned}$$

因为 A 有界, 故 $\inf_{z\in\overline{A}}\int_{G-z}g(y)\mathrm{d}y>0$. 而 $\lim_{x\to\infty}\int_{G-x}g(y)\mathrm{d}y=0$, 因此 $\lim_{x\to\infty}\mathbb{P}^x(T_A<\infty)=0$. 推出 A 暂留. □

Brown 运动实际上是经典的 Newton 位势理论的另一种表述. 下面我们给出经典的 Dirichlet 问题的概率解, 这是由 Kakutani(1944) 和 Doob(1954) 发现的.

定义 4.4.2　区域 $D\subset\mathbf{R}^d$ 上的函数 u 称为在 D 上调和, 如果 u 是局部可积且球面平均性质成立, 即对任何 $x\in D$, 存在 $\epsilon>0$, 当 $r<\epsilon$ 时有

$$u(x)=\int_{S_r(x)}u(y)\sigma_{x,r}(\mathrm{d}y),$$

其中 $\sigma_{x,r}$ 是球面 $S_r(x)$ 上的均匀分布.

一个已知的事实是 u 在 D 上调和当且仅当 u 无穷次可微且 $\Delta u=0$. 所谓的 Dirichlet 问题是这样的: 给定边界 ∂D 上的一个连续函数 f, 上面条件下存在一个在 $\overline{D}$ 上连续在 D 上调和的函数 u, 使得 $u|_{\partial D}=f$? 当然 Dirichlet 问题在分析中已

经有非常丰富深刻的结果. 我们在这里给出一个概率解只是为了说明 Brown 运动与位势之间的本质关系.

定理 4.4.4 设区域 $D \subset \mathbf{R}^d$ 是有界的且 D^c 是正则的, 则上述 Dirichlet 问题有唯一解

$$u(x) = \mathbb{E}^x f(B_T), \ x \in \overline{D}, \tag{4.4.1}$$

其中 T 是 D^c 的首中时.

证明 因 D 有界, 故 $\mathbb{P}^x(T < \infty) = 1$. 由连续性推出 $B_T \in \partial D$, 故等式 (4.4.1) 右边是有意义的.

首先唯一性是容易的. 事实上, 因 u 在有界闭集 $\overline{D}$ 上连续, 故达到最大最小值. 而调和函数球面平均性质推出 u 不可能在 D 上达到最大最小值, 只可能在边界上达到. 因此唯一性成立.

下面分三步证明 (4.4.1) 中定义的函数 u 是 Dirichlet 问题的解.

(1) 验证 u 在 ∂D 上是 f. 因为 D^c 是正则的, 故对 $x \in \partial D$, $u(x) = \mathbb{E}^x f(B_T) = \mathbb{E}^x f(B_0) = f(x)$.

(2) 验证 u 在 D 上调和. 对任何 $x \in D$, 存在 $\epsilon > 0$, 使得以 x 为中心, 以 ϵ 为半径的闭球包含在 D 中. 对任何 $0 < r < \epsilon$, 令

$$\tau := \inf\{t > 0 : |B_t - B_0| \geqslant r\}.$$

自然 $\mathbb{P}^x$ a.s., 有 $\tau < \infty$ 且 B_τ 在球面 $S_r(x)$ 上. 因为标准 Brown 运动是旋转不变的, 故当 $x = 0$ 时, B_τ 是 $S_r(0)$ 上的均匀分布. 又由平移不变性知一般地 B_τ 是 $S_r(x)$ 上的均匀分布.

直观地, 从 x 点出发, 在遇到 D^c 前一定会先遇到 $S_r(x)^c$, 即 P^x a.s., 有 $T \circ \theta_\tau + \tau = T$. (读者应该自行地严格证明之.) 由强 Markov 性,

$$\begin{aligned}\int_{S_r(x)} u(y)\sigma_{x,r}(\mathrm{d}y) &= \mathbb{E}^x u(B_\tau) = \mathbb{E}^x \mathbb{E}^{B_\tau} f(B_T) \\ &= \mathbb{E}^x f(B_T)\circ\theta_\tau = \mathbb{E}^x f(B(T\circ\theta_\tau + \tau)) = \mathbb{E}^x f(B_T) = u(x).\end{aligned}$$

因此 u 在 D 上调和.

(3) 验证 u 在 $\overline{D}$ 上连续. 只需验证在 ∂D 上连续. 任取 $a \in \partial D$, 要证明

$$\lim_{x\to a, x\in D} \mathbb{E}^x f(B_T) = f(a).$$

因为 f 是在 ∂D 上其实是一致连续的, 故只需验证对任何 $\epsilon > 0$, 有

$$\lim_{x\to a} \mathbb{P}^x(|B_T - a| > \epsilon) = 0.$$

而这等同于 $\lim_{x\to a} \mathbb{P}^x(|B_T - x| > \epsilon) = 0$. 任取 $t > 0$, 用 $\tau_\epsilon(x)$ 表示 Brown 运动从 x 出发首次离开半径为 ϵ 的球的时间,

$$\begin{aligned}\mathbb{P}^x(|B_T - x| > \epsilon) &= \mathbb{P}^x(|B_T - x| > \epsilon, T \leqslant t) + \mathbb{P}^x(|B_T - x| > \epsilon, T > t) \\ &\leqslant \mathbb{P}^x(\tau_\epsilon(x) \leqslant t) + \mathbb{P}^x(T > t) = \mathbb{P}(\tau_\epsilon(0) \leqslant t) + \mathbb{P}^x(T > t).\end{aligned}$$

显然 $\mathbb{P}(\tau_\epsilon(0) = 0) = 0$, $\mathbb{P}^a(T > 0) = 0$. 对任何固定的 $t > 0$, 只要 $x \mapsto \mathbb{P}^x(T > t)$ 是上半连续的, 便推出 $\limsup_{x\to a} \mathbb{P}^x(T > t) \leqslant \mathbb{P}^a(T > t) = 0$, 即 $\lim_{x\to a} \mathbb{P}^x(T > t) = 0$. 因此在上面不等式两边取极限得

$$\lim_{x\to a} \mathbb{P}^x(|B_T - x| > \epsilon) \leqslant \mathbb{P}(\tau_\epsilon(0) \leqslant t)).$$

而右边是个无穷小量, 推出所需结论. 故最后我们需要证明 $x \mapsto \mathbb{P}^x(T > t)$ 上半连续.

事实上, 对任何 $0 < s < t$, $\mathbb{P}^x(T\circ\theta_s + s \leqslant t) = \mathbb{E}^x(\mathbb{P}^{B_s}(T \leqslant t - s))$ 是个关于 x 的连续函数. 而 $\mathbb{P}^x(T \leqslant t)$ 是 $\mathbb{P}^x(T\circ\theta_s + s \leqslant t)$ 的递增极限, 故 $\mathbb{P}^\cdot(T \leqslant t)$ 是下半连续的, 即 $\mathbb{P}^\cdot(T > t)$ 上半连续. □

习 题

1. 以下习题中, 设 $B = (B_t)_{t\geqslant 0}$ 是 1-维标准 Brown 运动. 设 $a > 0$, $t \geqslant 0$, 定义

$$B_t^a = \begin{cases} B_t, & t < T_a; \\ 2a - B_t, & t \geqslant T_a. \end{cases}$$

   证明: B^a 等价于 B.

2. 设 $b \geqslant a > 0$, $S_t := \sup_{s\in[0,t]} B_s$. 证明:

$$\mathbb{P}(S_t > b, B_t < a) = \mathbb{P}(B_t < a - 2b) = \mathbb{P}^{2b}(B_t < a),$$

并求 (B_t, S_t) 的联合密度.

3. 证明: $S_t - B_t$ 与 $|B_t|$ 的分布相同.
4. 证明: (B_t, S_t) 是状态空间为 $\{(x,y) : x \leqslant y, y > 0\}$ 的 Markov 过程, 并计算其转移函数.
5. 设 $T = T_{(-\infty,0]}$, 那么过程 $B_{t\wedge T}$ 称为是 $\mathbf{R}_+$ 上具有吸收壁 0 的 Brown 运动, 而过程 $|B|$ 也是右过程, 称为是 $\mathbf{R}_+$ 上具有反射壁 0 的 Brown 运动. 分别计算它们的转移函数.
6. 证明: 过程 $\{T_a : a \geqslant 0\}$ 是左连续的, 有平稳独立增量在任何区间上都不连续的增过程. 如果定义 $T_{a+} := \lim_{b\downarrow a} T_b$, 证明: 对任何 $a \geqslant 0$, $T_a = T_{a+}$ a.s..
7. 证明:

$$\lim_{t\to+\infty} \sqrt{t}\mathbb{P}(B_s \leqslant 1, \forall s \leqslant t) = \sqrt{\frac{2}{\pi}}.$$

8. 证明: 集合 A 的正则点一定是 A 的 Euclid 意义下的极限点. 对 1-维 Brown 运动, 反之亦然.
9. 设 T 是集合 A 的首中时. 证明: 对任何 $t \geqslant 0$, 当 $T > t$ 时有 $T\circ\theta_t + t = T$. 由此证明定理 4.4.4 中的用到的一个断言: $T\circ\theta_\tau + \tau = T$ a.s..
10. 证明: 一维 Brown 运动的精细拓扑与通常拓扑一致. 而二维以上的 Brown 运动的精细拓扑严格细于通常拓扑.

§4.5 局部时与游离理论

下面介绍 Brown 运动的局部时及其性质, 这个概念的引入要归功于 Lévy. 在 §3.5 中我们介绍了一个连续半鞅的局部时. 当然关于 Brown 运动的局部时是首先被引入的也是最多被讨论的. 现在设 B 是 Brown 运动, 对任何 $x \in \mathbf{R}$, 存在一个连续增过程 L_t^x, 称为 Brown 运动在 x 点的局部时, 使得对任何 t,

$$(B_t - x)^+ - (B_0 - x)^+ = \int_0^t 1_{(0,\infty)}(B_s - x)\mathrm{d}B_s + \frac{1}{2}L_t^x;$$

$$|B_t - x| - |B_0 - x| = \int_0^t \mathrm{sgn}(B_s - x)\mathrm{d}B_s + L_t^x.$$

特别地, 写 $L_t = L_t^0$, Brown 运动在 0 点的局部时. 下面是局部时的直观意义.

引理 4.5.1　对任何 $t > 0$, 在 L^2 收敛的意义下有

$$L_t = \lim_{\delta \downarrow 0} \frac{1}{2\delta} \int_0^t 1_{(-\delta,+\delta)}(B_s)\mathrm{d}s.$$

上式在几乎处处意义下对所有 t 成立.

证明　作函数 f_δ 满足 $f_\delta(-\infty) = 0$ 且

$$f_\delta'' = \frac{1}{2\delta} 1_{(-\delta,+\delta)},$$

是个凸函数. 取 $g \in C_c^\infty$ 且 $\int g\mathrm{d}x = 1$, 令 $g_n(x) = ng(nx)$, $f_n = g_n * f_\delta$, 那么 $f_n \to f_\delta$, $f_n' \to f_\delta'$. 而除了两个点 $\{-\delta, +\delta\}$ 外, $f_n'' \to f_\delta''$. 由 Itô 公式,

$$f_n(B_t) - f_n(B_0) = \int_0^t f_n'(B_s)\mathrm{d}B_s + \frac{1}{2}\int_0^t f_n''(B_s)\mathrm{d}s.$$

让 $n \to \infty$, 因为 B_s 在 $[0,t]$ 上等于两个固定点的 s 集合的测度为零, 因此几乎处处地

$$f_\delta(B_t) - f_\delta(B_0) = \int_0^t f_\delta'(B_s)\mathrm{d}B_s + \frac{1}{2}\int_0^t f_\delta''(B_s)\mathrm{d}s.$$

而当 $\delta \to 0$ 时, 因为 $|f_\delta(B_t) - f_\delta(B_0)| \leqslant |B_t - B_0|$, 故上式左边 L^2 收敛于 $B_t^+ - B_0^+$. 右边第一项也 L^2 收敛, 因为由随机积分性质

$$\begin{aligned}
&\mathbb{E}\left(\int_0^t f_\delta'(B_s) - 1_{(0,\infty)}(B_s)\mathrm{d}B_s\right)^2 \\
&\leqslant \mathbb{E}\int_0^t 1_{[-\delta,+\delta]}(B_s)\mathrm{d}s \\
&= \int_0^t \mathbb{P}(|B_s| \leqslant \delta)\mathrm{d}s \leqslant \int_0^t \frac{2\delta}{\sqrt{2\pi s}}\mathrm{d}s \to 0.
\end{aligned}$$

故而右边第二项, 即引理中的右边 L^2 收敛. □

下面证明 $\mathrm{d}L_t$ 的支撑 $\operatorname{supp} L$ 恰好是 $Z = \{t : B_t = 0\}$. 首先由 Tanaka 公式 $\operatorname{supp} L \subset Z$. 要证明 $\operatorname{supp} L \supset Z$, 需要一个引理.

引理 4.5.2 (Skorohod) 设 y 是 $[0,\infty)$ 上实值连续函数且 $y(0) \geqslant 0$, 则存在唯一的 $[0,\infty)$ 上连续函数对 (x,z) 满足:

(1) $z = x + y$;

(2) z 是非负的;

(3) x 是递增连续函数, $x(0) = 0$ 且 $\mathrm{d}x(t)$ 支撑在 $\{t : z(t) = 0\}$ 上.

实际上, $x(t) = \sup_{s\leqslant t}(-y(s))^+ = -\inf_{s\leqslant t} y(s)^-$.

证明 验证 $x(t) := \sup_{s\leqslant t}(-y(s))^+$ 和 $z := x + y$ 满足条件 (1), (2), (3) 的工作读者可自行完成. 关键是证明唯一性. 设 $(\hat{x},\hat{z})$ 也满足条件, 那么 $z - \hat{z} = x - \hat{x}$, 因 (3), 故

$$
\begin{aligned}
(z(t) - \hat{z}(t))^2 &= 2\int_0^t (z(s) - \hat{z}(s))\mathrm{d}(x(s) - \hat{x}(s)) \\
&= -2\int_0^t \hat{z}(s)\mathrm{d}x(s) - 2\int_0^t z(s)d\hat{x}(s) \leqslant 0.
\end{aligned}
$$

因此 $z = \hat{z}$, $x = \hat{x}$. □

由局部时的性质, $|B_t| = \beta_t + L_t$, 其中 $\beta_t = \int_0^t \operatorname{sgn}(B)\mathrm{d}B$. 显然 β_t 是连续鞅且 $\langle\beta\rangle_t = t$, 由 Lévy 刻画定理 3.4.12, β 也是一个标准 Brown 运动. 因此由上面的引理推出

$$
L_t = -\inf_{s\leqslant t} \beta_s^-.
$$

这得出两个推论:

(1) 对任何 t, $L_t > 0$ a.s., 即 $0 \in \operatorname{supp} L$;

(2) $\lim\limits_{t\to\infty} L_t = \infty$.

定理 4.5.1 局部时在且只在 Brown 运动的零点增加, 精确地 $\operatorname{supp} L = Z$ a.s..

证明 只需验证 $\operatorname{supp} L \supset Z$. 对任何 t, 令 $r_t := \inf\{s > t : B_s = 0\}$, 则 r_t 是停时且由强 Markov 性 $s \mapsto B_{r_t+s}$ 是标准 Brown 运动. 它的局部时是 $s \mapsto L_{r_t+s} - L_{r_t}$. 故对任何 $s > 0$, $L_{r_t+s} > L_{r_t}$ a.s. 推出 $r_t \in \operatorname{supp} L$ a.s., 或对几乎所有 ω, 对所有有理

数 t, $r_t(\omega) \in \operatorname{supp} L(\omega)$. 现在在这样的 ω 上考虑, 如果 $a \notin \operatorname{supp} L$, 存在 a 的领域 $U_a \cap \operatorname{supp} L = \varnothing$. 取有理数 $t < a$ 且 $a \in U_a$, 则 $t \leqslant r_t$ 且 $r_t \in \operatorname{supp} L$. 故 $r_t \notin U_a$ 推出 $a \notin Z$. 因此有 $Z \subset \operatorname{supp} L$. □

设 (τ_t) 是 L_t 的右连续逆, 即 $\tau_t := \inf\{s > 0 : L_s > t\}$, 那么 τ_t 是停时且 $L_{\tau_t} = t$, $\tau_{L_t} = r_t$. 因 $\tau t- < \tau_t$ 当且仅当 L 在 $[\tau_{t-}, \tau_t]$ 上是常数, 而 $\operatorname{supp} L$ 实际上是 $\mathrm{d}L_t$ 测度为零的最大开集的补集, 故

$$(\operatorname{supp} L)^c = \bigcup_{t>0} (\tau_{t-}, \tau_t) = Z^c.$$

引理 4.5.3 (1) 设 T 是停时, $t > 0$ 是正随机变量, 则 $L_{T+t} = L_T + L_t \circ \theta_T$;

(2) 对任何 $t, s > 0$, $\tau_{t+s} = \tau_t + \tau_s \circ \theta_{\tau_t}$.

证明　利用引理 4.5.1, 得

$$\begin{aligned}
L_{T+t} &= \lim_{\delta \to 0} \frac{1}{2\delta} \int_0^{T+t} 1_{(-\delta,+\delta)}(B_s)\mathrm{d}s \\
&= L_T + \lim_{\delta \to 0} \frac{1}{2\delta} \int_T^{T+t} 1_{(-\delta,+\delta)}(B_s)\mathrm{d}s \\
&= L_T + \lim_{\delta \to 0} \frac{1}{2\delta} \int_0^{t} 1_{(-\delta,+\delta)}(B_s \circ \theta_T)\mathrm{d}s \\
&= L_T + L_s \circ \theta_T.
\end{aligned}$$

利用上面的结论计算

$$\begin{aligned}
\tau_{t+s} &= \inf\{r > 0 : L_r > t + s\} = \inf\{r > 0 : r > \tau_t, L_r > t + s\} \\
&= \inf\{r > 0 : L_{r+\tau_t} - L_{\tau_t} > s\} + \tau_t \\
&= \inf\{r > 0 : L_r \circ \theta_{\tau_t} > s\} + \tau_t \\
&= \tau_s \circ \theta_{\tau_t} + \tau_t.
\end{aligned}$$

□

现在我们可以简单地介绍由 Itô 开创的 Brown 运动的游离(excursion)理论. 因为 Brown 运动的零点集是无处稠零测度的完备集, 故其离开 0 点的时间是一个稠密开集, 或者说是开区间的并, 称为是 Brown 运动关于零点的游离. (其实对其他任

何点都是一样的.) 现在设 $B=(B_t,\mathbb{P})$ 是 0 点出发的标准 Brown 运动. 不妨设样本空间 W 是 0 点出发的连续函数全体. 对 $w\in W$, 令 $R(w):=\inf\{t>0:w(t)=0\}$. E 是 W 中满足 $R(w)>0$, 且对任何 $t>R(w)$ 有 $w(t)=0$ 的连续函数 w 全体. 注意, $\mathbb{P}(U)=0$ 因为 $\mathbb{P}(R=0)=1$. δ 是恒等于 0 的连续函数. 对 $w\in W$, $s>0$, 如果 $\tau_s(w)>\tau_{s-}(w)$, 那么 $(\tau_{s-}(w),\tau_s(w))$ 是 w 游离 0 点的一个极大区间, 定义

$$e_s(w)=\{w(t+\tau_{s-})1_{(0,\tau_s(w)-\tau_{s-}(w))}:\ t\in[0,\infty)\};$$

右边是 U 中的一个函数. 如果 $\tau_s(w)=\tau_{s-}(w)$, 定义 $e_s:=\delta$. 因为 τ 的不连续点可数, 因此 (e_s) 是以 U_δ 为状态空间的点过程. 它记录了 Brown 运动的所有游离区间, 称为游离过程.

首先 (e_s) 是 σ- 离散的. 令 $U_n:=\{w\in U:\ R(w)>\dfrac{1}{n}\}$, 那么 $U_n\uparrow U$ 且若 $N_e((0,t]\times U_n)=\#\{s\leqslant t:\tau_s-\tau_{s-}>\dfrac{1}{n}\}$, 则 $\tau_t\geqslant\dfrac{1}{n}N_e((0,t]\times U_n)$, 因此 $N_e((0,t]\times U_n)\leqslant n\tau_t<\infty$ a.s..

定理 4.5.2 (Itô) 游离过程 (e_s) 是一个平稳 $(\mathscr{F}_{\tau_t})$-Poisson 点过程.

证明 利用定理 2.5.5 证明之. 首先上面引理推出 $e_{s+r}=e_r\circ\theta_{\tau_s}$. 因此对任何 $s>0,0<t_1<\cdots<t_d$ 及 $U_i\in U$, 由强 Markov 性推出

$$\begin{aligned}
&\mathbb{P}(N_e((s,s+t_i]\times U_i):1\leqslant i\leqslant d|\mathscr{F}_{\tau_s})\\
&=\mathbb{P}(N_e((0,t_i]\times U_u)\circ\theta_{\tau_s}:1\leqslant i\leqslant d|\mathscr{F}_{\tau_s})\\
&=\mathbb{E}(\mathbb{P}^{B_{\tau_s}}(N_e((0,t_i]\times U_u):1\leqslant i\leqslant d)\\
&=\mathbb{P}(N_e((0,t_i]\times U_u):1\leqslant i\leqslant d).
\end{aligned}$$

完成证明. □

Itô 还给出了游离过程在 U 上的特征测度 n, 它也称为是 Itô 测度. 在这里我们不加证明地叙述这个漂亮的结果, 有兴趣的读者可参考文献 [29]. 设 (Q_t) 是 B 遇到 0 时停止的过程的半群, 即 $B_{t\wedge T}$ 的转移半群, 其中 T 是 0 点的首中时, 那么

$$Q_t(x,\mathrm{d}y)=\frac{1}{\sqrt{2\pi t}}\left(\mathrm{e}^{-\frac{(y-x)^2}{2t}}-\mathrm{e}^{-\frac{(y+x)^2}{2t}}\right)1_{\{xy>0\}}\mathrm{d}y.$$

令

$$\mu_t(\mathrm{d}y) := \frac{1}{\sqrt{2\pi t^3}}|y|\mathrm{e}^{-\frac{y^2}{2t}}\mathrm{d}y,$$

那么 n 的有限维分布如下: $0 < t_1 < t_2 < \cdots < t_k$, $y_1 \neq 0, \cdots, y_k \neq 0$,

$$\begin{aligned} n(\{w \in U : & w(t_1) \in \mathrm{d}y_1, \cdots, w(t_k) \in \mathrm{d}y_k\}) \\ &= \mu_{t_1}(\mathrm{d}y_1)Q_{t_2-t_1}(y_1, \mathrm{d}y_2)\cdots Q_{t_k-t_{k-1}}(y_{k-1}, \mathrm{d}y_k). \end{aligned}$$

局部时实际上是两个参数的随机过程 $(t,x) \mapsto L_t^x(\omega)$. 已知固定 x 关于变量 t 是连续的, 下面我们证明它有一个修正是关于 (t,x) 联合连续的. 首先

$$\frac{1}{2}L_t^x = (B_t - x)^+ - (B_0 - x)^+ - \int_0^t 1_{(x,\infty)}(B_s)\mathrm{d}B_s.$$

因此只需验证 $G(t,x) := \int_0^t 1_{(x,\infty)}(B_s)\mathrm{d}B_s$ 有一个连续修正. 类似于 Brown 运动构造, 只需证明对任何 $T > 0$, 存在常数 $C > 0$, 使得对任何 $x, y \in \mathbf{R}$,

$$\mathbb{E} \sup_{t\in[0,T]} |G(t,x) - G(t,y)|^4 \leqslant C(x-y)^2.$$

因 $G(\cdot,x) - G(\cdot,y)$ 是连续鞅, 故由 Doob 鞅不等式,

$$\mathbb{E} \sup_{t\in[0,T]} |G(t,x) - G(t,y)|^4 \leqslant \left(\frac{4}{3}\right)^4 \mathbb{E}|G(T,x) - G(T,y)|^4.$$

要估计右边, 需要一个不等式: 对任何连续鞅 M, 有

$$\mathbb{E}M_t^4 \leqslant 36\mathbb{E}\langle M\rangle_t^2.$$

让我们先承认这个不等式, 那么

$$\begin{aligned} \mathbb{E}|G(T,x) - G(T,y)|^4 &\leqslant 36\mathbb{E}\left(\int_0^T 1_{(x,y]}(B_s)\mathrm{d}s\right)^2 \\ &= 36\int_0^T\int_0^T \mathbb{P}(B_s \in (x,y])\mathbb{P}(B_t \in (x,y])\mathrm{d}s\mathrm{d}t \\ &\leqslant 36(x-y)^2\int_0^T\int_0^T \frac{1}{\sqrt{2\pi t}}\frac{1}{\sqrt{2\pi s}}\mathrm{d}t\mathrm{d}s \\ &\leqslant \frac{72T}{\pi}(x-y)^2. \end{aligned}$$

上面的鞅不等式证明如下. 由 Itô 公式,

$$\begin{aligned}\mathbb{E}M_t^4 &= 6\mathbb{E}\int_0^t M_s^2 \mathrm{d}\langle M\rangle \\ &= 6\mathbb{E}M_t^2\langle M\rangle_t - 6\mathbb{E}\int_0^t \langle M\rangle_s \mathrm{d}M_s^2 \\ &= 6\mathbb{E}M_t^2\langle M\rangle_t - 6\mathbb{E}\int_0^t \langle M\rangle_s \mathrm{d}\langle M\rangle_s \\ &\leqslant 6\mathbb{E}M_t^2\langle M\rangle_t,\end{aligned}$$

然后应用 Cauchy-Schwarz 不等式.

因此不妨设 $(t,x)\mapsto L_t^x$ 是连续的.

定理 4.5.3 (Itô-Tanaka) 设 f 是凸函数, 则其二阶导数是一个 Radon 测度 μ_f,

$$f(B_t) = f(B_0) + \int_0^t f'_-(B_s)\mathrm{d}B_s + \frac{1}{2}\int_{\mathbf{R}} L_t^x \mu_f(\mathrm{d}x),$$

其中 μ_f 是 f 的 Schwarz 意义下的二阶导数.

证明 首先容易验证随机积分的 Fubini 定理: 如果 F 是 $\mathbf{R}^2$ 上有界可测函数, μ 是 $\mathbf{R}$ 上有限测度, 那么

$$\int_{\mathbf{R}} \mathrm{d}x \int_0^t F(x,B_s)\mathrm{d}B_s = \int_0^t \mathrm{d}B_s \int_{\mathbf{R}} F(x,B_s)\mathrm{d}x.$$

不妨设 μ_f 是紧支撑的, 那么存在 a,b, 使得

$$f(y) := a + by + \frac{1}{2}\int_{\mathbf{R}} |y-x|\mu_f(\mathrm{d}x).$$

故 $f'_-(y) = b + \int_{\mathbf{R}} \frac{1}{2}\mathrm{sgn}(y-x)\mu_f(\mathrm{d}x)$. 对 Tanaka 公式两边关于 μ_f 积分并应用 Fubini 定理便推出定理结论. □

设 g 是紧支撑连续函数, $\mu_f(\mathrm{d}x) = g(x)\mathrm{d}x$, 那么比较 Itô 公式推出

$$\int_0^t g(B_s)\mathrm{d}s = \int_{\mathbf{R}} L_t^x g(x)\mathrm{d}x.$$

此公式对于 Borel 可测函数 g 仍然成立. 称为占据时间公式.

习　　题

1. 设 $S_t := \sup_{s\leqslant t} B_s$, 证明: 过程 $(S_t - B_t, S_t)$ 与 $(|B_t|, L_t)$ 等价.
2. 证明 Itô-Tanaka 定理证明中所叙述的关于随机积分的 Fubini 定理.
3. 设 $f(x) = |x|$, 证明:

$$L_t = \lim_{s\to 0} \frac{1}{s}\int_0^t (P_s f(B_u) - f(B_u))\mathrm{d}u.$$

4. 证明: 局部时的逆 (τ_t) 是两个严格增的 Lévy 过程.
5. 设 $B = (B_t^1, B_t^2)$ 是标准 2 维 Brown 运动. τ_t 是 B^1 的局部时的逆, 证明: $B_{\tau_t}^2$ 是对称 Cauchy 过程.

§4.6　Markov 过程的变换

设

$$X = (\Omega, \mathscr{F}, (\mathscr{F}_t), (X_t), (\theta_t), \mathbb{P}^x)$$

是状态空间 E 上的右 Markov 过程, 其半群和豫解分别是 (P_t), (U^α). 我们先来介绍乘泛函.

定义 4.6.1　一个实值右连续随机过程 $M = (M_t : t \geqslant 0)$ 称为是 X 的乘泛函, 如果下列条件满足:

(1) 对任何 $t \geqslant 0$, 有 $M_t \in [0,1]$;

(2) $M = (M_t)$ 是 $(\mathscr{F}_t)$ 适应的;

(3) 对几乎所有的 $\omega \in \Omega$, 有

$$M_{t+s}(\omega) = M_s(\omega)\cdot M_t(\theta_s\omega)$$

对任何 $t, s \geqslant 0$ 成立.

一个乘泛函 M 称为是恰好的, 如果对任何 $t > 0$, $t_n \downarrow\downarrow 0$, 有 $M_{t-t_n}\circ\theta_{t_n} \to M_t$.

条件 (3) 可以由看上去更弱的条件 (3′) 代替:

(3′) 对任何 $t, s \geqslant 0$, $M_{t+s} = M_s \cdot M_t \circ \theta_s$ a.s..

由 (3′) 推出 (3) 的证明过程称为 M 的完美化, 参考文献 [30], [15]. 上面的条件 (1),(3) 蕴含着 M 是递减的, 故这样的过程也通常称为递减乘泛函, 如果把条件 (2) 去掉, 那么 M 称为原始的乘泛函.

例 4.6.1　对任何 $\alpha > 0$, $t \mapsto \mathrm{e}^{-\alpha t}$ 总是乘泛函. 设 T 是随机时间, 令 $M_t := 1_{\{t<T\}}$, 那么 M 是满足 (1) 的右连续随机过程. 容易验证 M 适应当且仅当 T 是停时; M 满足 (3) 当且仅当 T 是原始的终止时, 还应该注意末离时不是停时, 但是一个原始的终止时; 首中时是恰好的而进入时不是恰好的. ∎

设 M 是乘泛函. 由乘性 (3) 知, $M_0^2 = M_0$, 再由 Blumenthal 0-1 律推出对任何 $x \in E$, $\mathbb{P}^x(M_0 = 1) = 0$ 或者 1. 定义

$$E_M := \{x \in E:\ \mathbb{P}^x(M_0 = 1) = 1\},$$

$$S_M := \inf\{t > 0:\ M_t = 0\},$$

它们分别称为是 M 的永久点集与生命时. 说 M 是右的乘泛函, 如果 E_M 是近乎可选的. 随机变量 S_M 是停时, 因为 $\{S_M \leqslant t\} = \{M_t = 0\} \in \mathscr{F}_t$. 另外 S_M 还是终止时, 因为如果 $S_M > t$, 即 $M_t > 0$, 那么

$$\begin{aligned} S_M \circ \theta_t &= \inf\{s > 0:\ M_s \circ \theta_t = 0\} \\ &= \inf\{s > 0:\ M_{t+s} = 0\} = S_M - t. \end{aligned}$$

当 $S_M \geqslant \zeta$ 时, 我们说 M 恒正.

对 E 上非负 $\mathscr{E}^*$-可测函数 f, 定义

$$Q_t f(x) := \mathbb{E}^x(f(X_t) M_t),\ x \in E,\ t \geqslant 0.$$

引理 4.6.1　(Q_t) 是 E 上的转移半群.

证明　因为 $f(X_t)M_t$ 是 $\mathscr{F}_t$ 可测的, 故 $Q_t f \in \mathscr{E}^*$. 利用 Markov 性与 M 的乘性,

$$\begin{aligned} Q_t Q_s f(x) &= \mathbb{E}^x(Q_s f(X_t) M_t) \\ &= \mathbb{E}^x((f(X_s) M_s) \circ \theta_t M_t) \\ &= \mathbb{E}^x(f(X_{s+t}) \cdot M_{s+t}) = Q_{t+s} f(x). \end{aligned}$$

最后 $Q_t 1(x) = \mathbb{E}^x(M_t; X_t \in E) \leqslant 1$. □

两个乘泛函的乘积当然还是乘泛函, 乘泛函 M 与 $M1_{[0,\zeta)}$ 产生的半群是一样的, 因此我们不妨总假设 $S_M \leqslant \zeta$. 因为 (Q_t) 是 Markov 转移半群, 故应该有一个 Markov 过程对应于它或者说对应于 M, 怎么描述这个过程呢? 先定义 Killing 算子族 (k_t): 对任何 $t \geqslant 0$,

$$k_t\omega(s) := \begin{cases} \omega(s), & s < t; \\ \Delta, & s \geqslant t. \end{cases}$$

k_t 是 $(\Omega, \mathscr{F}^0)$ 上的可测映射, 它把轨道从 t 时刻截断放入坟墓. 显然 $X_s \circ k_t = X_s$, $s < t$, 而 $X_s \circ k_t = \Delta$, $s \geqslant t$. 对 $(\Omega, \mathscr{F}^{0*})$ 上的非负或有界随机变量 Z, 定义

$$\mathbb{Q}^x(Z) := \mathbb{E}^x \int_{(0,+\infty]} Z \circ k_t \mathrm{d}(-M_t),$$

其中 $M_\infty := 0$. 我们来证明

$$Q_t f(x) = \mathbb{Q}^x f(X_t),\ x \in E,\ t \geqslant 0.$$

由定义

$$\begin{aligned} \mathbb{Q}^x f(X_t) &= \mathbb{E}^x \int_{(0,\infty]} f(X_t) \circ k_u \mathrm{d}(-M_u) \\ &= \mathbb{E}^x \int_{(t,\infty]} f(X_t) \mathrm{d}(-M_u) \\ &= \mathbb{E}^x(f(X_t)M_t) = Q_t f(x). \end{aligned}$$

因此下面的定理是直观的.

定理 4.6.1 如果 M 是右的乘泛函, 那么过程

$$Y = (\Omega, \mathscr{F}, \mathscr{F}_t, X_t, \theta_t, \mathbb{Q}^x)$$

是 E_M 上, 转移半群为 (Q_t) 的右 Markov 过程, 其中强化是相对于概率 $(\mathbb{Q}^x : x \in E_M)$ 的.

证明 不难从定义直接验证它是一个右连续简单 Markov 过程. 关键是证明它满足

(HD2). 这证明涉及许多细节, 不适合在这里叙述, 可参考文献 [30] 中的 §61. 但如果 M 是恰好的, 证明要简单得多. 用 (V^α) 表示 (Q_t) 的豫解算子, 对 $\mathscr{E}^*$ 非负可测函数 f 和 $\alpha \geqslant 0$, 定义

$$P_M^\alpha f(x) = \begin{cases} \mathbb{E}^x \int_0^\infty \mathrm{e}^{-\alpha t} f(X_t) \mathrm{d}(-M_t), & x \in E_M, \\ f(x), & x \notin E_M. \end{cases}$$

当 $M = 1_{[0,T)}$ 时, $P_M^\alpha f = P_T^\alpha f$. 我们来证明关于算子 P_M^α 的几个事实. 首先

$$P_M^\alpha U^\alpha f = \mathbb{E}^x \int_0^\infty \mathrm{e}^{-\alpha s} f(X_s)(1 - M_s) \mathrm{d}s = U^\alpha f - V^\alpha f.$$

这实际上是豫解方程的推广. 事实上, 当 $x \notin E_M$ 时是显然的, 如果 $x \in E_M$, 那么利用 Markov 性和 Fubini 定理,

$$\begin{aligned} P_M^\alpha U^\alpha f(x) &= \mathbb{E}^x \int_0^\infty \mathrm{e}^{-\alpha t} \mathrm{d}(-M_t) \int_0^\infty \mathrm{e}^{-\alpha s} f(X_s) \circ \theta_t \mathrm{d}s \\ &= \mathbb{E}^x \int_0^\infty \mathrm{d}(-M_t) \int_{[t,\infty)} \mathrm{e}^{-\alpha s} f(X_s) \mathrm{d}s \\ &= \mathbb{E}^x \int_0^\infty \mathrm{e}^{-\alpha s} f(X_s) \mathrm{d}s \int_{(0,s]} \mathrm{d}(-M_t) \\ &= \mathbb{E}^x \int_0^\infty \mathrm{e}^{-\alpha s} f(X_s)(1 - M_s) \mathrm{d}s. \end{aligned}$$

注意第一个等号要用到一个重要公式, 见文献 [30] 中 (32.6). 下面再证明 $P_M^\alpha U^\alpha f$ 是 α-上平均的, 且当 M 是恰当时, 它是 α-过分的. 事实上, 由 Markov 性, 且因为 $1 - M_s \circ \theta_t \leqslant 1 - M_{s+t}$, 故

$$\begin{aligned} P_t^\alpha P_M^\alpha U^\alpha f(x) &= \mathbb{E}^x \int_0^\infty \mathrm{e}^{-\alpha(s+t)} f(X_{s+t})(1 - M_s \circ \theta_t) \mathrm{d}s \\ &\leqslant \mathbb{E}^x \int_0^\infty \mathrm{e}^{-\alpha(s+t)} f(X_{s+t})(1 - M_{s+t}) \mathrm{d}s \\ &= \mathbb{E}^x \int_t^\infty \mathrm{e}^{-\alpha s} f(X_s)(1 - M_s) \mathrm{d}s \end{aligned}$$

$$\leqslant P_M^\alpha U^\alpha f(x).$$

这是 α-上平均性. 做变量替换, 我们有

$$P_t^\alpha P_M^\alpha U^\alpha f(x) = \mathbb{E}^x \int_t^\infty \mathrm{e}^{-\alpha s} f(X_s)(1 - M_{s-t}\circ\theta_t)\mathrm{d}s.$$

如果 M 是恰好的, 那么当 $t \downarrow\downarrow 0$ 时, $M_{s-t}\circ\theta_t \downarrow M_s$, 因此由单调收敛定理推出

$$P_t^\alpha P_M^\alpha U^\alpha f(x) \uparrow P_M^\alpha U^\alpha f(x).$$

这样我们证明了如果 $f \in C_u(E)$, 那么 $V^\alpha f$ 是两个 α- 过分函数的差, 故由定理 4.2.3 推出 $V^\alpha f\circ X.$ 是右连续的. □

由定理 4.2.9 看出来如果 M 是恰好的, 那么 $V^\alpha f$ 是精细连续函数的差, 因此也是精细连续的, 故 $E_M = \{V^\alpha 1 > 0\}$ 是精细开集.

X 经过其一个乘泛函 M 而得到过程 Y 的方法称为是 Markov 过程的 Killing 变换, 它是 Markov 过程理论中常见变换之一, Y 称为是 X 的 M-子过程. 如果 $M_t = \mathrm{e}^{-\alpha t}$, M-子过程也称为 α-子过程. 如果 G 是 E 的开子集, τ 是 G^c 的进入时, 那么 $M = 1_{[0,\tau)}$ 的永久点集就是 G, 因此 M 是右的, M-子过程也称为 X 在 G 上的限制. 一般地, M-子过程的半群 (Q_t) 被半群 (P_t) 控制, 即对任何非负可测函数 f, 和 $t \geqslant 0$, 有 $Q_t f \leqslant P_t f$. 下面定理说明反过来也成立, 证明参考文献 [5].

定理 4.6.2 设有典则轨道空间上以 E 为状态空间的两个右过程 $X = \{\mathbb{P}^x\}$ 和 $Y = \{\mathbb{Q}^x\}$, 则 Y 的转移半群 (Q_t) 被 X 的转移半群 (P_t) 控制当且仅当存在 X 的乘泛函 M, 使得 Y 是 X 的 M-子过程.

如果把定义 4.6.1 的条件 (1) 改为 (1′): 对任何 $x \in E$, $t \geqslant 0$ 有 $M_t \geqslant 0$ 且 $\mathbb{E}^x M_t \leqslant 1$, 那么 M 是一个非负上鞅, 称为上鞅乘泛函. 进一步, 如果 $\mathbb{E}^x M_t = 1$, 那么 M 是一个非负鞅, 称为鞅乘泛函.

例 4.6.2 设 $h \in \mathbf{S}^\alpha$, 那么 $\mathbb{E}^x \mathrm{e}^{-\alpha t} h(X_t) \leqslant h(x)$. 令

$$M_t := \mathrm{e}^{-\alpha t}\frac{h(X_t)}{h(X_0)},\ E_h := \{0 < h < +\infty\},$$

那么 M 是上鞅乘泛函且 $E_M = E_h$.

给定上鞅乘泛函 M, 类似地定义半群 (Q_t), 那么它也对应一个右过程, 这也是

Markov 过程中一类重要的变换, 参考文献 [30], §62. 上面例子中这样的 h 诱导的变换称为是 h-变换, 是由 Doob 引入的. Itô-Watanabe 分解告诉我们上鞅乘泛函总可以表示为鞅乘泛函和连续递减乘泛函的乘积. 鞅乘泛函诱导的变换称为是漂移变换. 下面我们来介绍加泛函和时间变换, 它的定义与乘泛函类似.

定义 4.6.2 设 M 是 X 的乘泛函. 一个实值右连续随机过程 $A = (A_t : t \geqslant 0)$ 称为是 X 的 M-加泛函, 如果下列条件满足:

(1) 对任何 $t < S_M$, 有 $0 \leqslant A_t < +\infty$;

(2) $A = (A_t)$ 是 $(\mathscr{F}_t)$ 适应的;

(3) 对几乎所有的 $\omega \in \Omega$, 有

$$A_{t+s}(\omega) = A_s(\omega) + M_s(\omega) \cdot A_t(\theta_s \omega)$$

对任何 $t, s \geqslant 0$ 成立.

特别地, $1_{[0,\zeta)}$-加泛函简称为加泛函.

如同乘泛函, 加泛函也可以完美化. 非负性和加性推出 A 一定是增过程. 另外由 (3) 推出如果 $x \in E_M$, 那么 $\mathbb{P}^x(A_0 = 0) = 1$. 条件 (3) 也简单写为 $A_{t+s} = A_s + M_s \cdot A_t \circ \theta_s$. 由此推出 A_t 在 $\{t > S_M\}$ 上是常数. 如果没有适应性(2), 加泛函称为原始的加泛函. 如果 A 是加泛函且 $t \mapsto A_t$ 是连续的, 那么 A 称为连续加泛函.

例 4.6.3 如果 M 是乘泛函, 那么 $1 - M_{s+t} = 1 - M_s + M_s - M_s \cdot M_t \circ \theta_s = (1 - M_s) + M_s \cdot (1 - M_t) \circ \theta_s$, 因此 $(1 - M_t : t \geqslant 0)$ 总是 M-加泛函. 如果 f 是非负有界 $\mathscr{E}^*$ 可测函数, 那么

$$A_t := \int_0^t f(X_s) \mathrm{d}s, \ t \geqslant 0$$

定义了一个连续加泛函. 特别地, $A_t = \alpha t$ 总是连续加泛函, 其中 α 是非负常数. 如果 A 是一个加泛函, 那么容易验证 $\exp(-A_t)$ 是一个乘泛函. 反过来, 如果 M 是乘泛函, 其对数未必是加泛函, 但若 M 恒正, 那么容易验证其自然对数 $\log M_t^{-1}$ 定义了一个加泛函. 但自然对数并不是最好的选择, 我们定义 M 的 Stieltjes 对数

$$(\log M)_t := \int_0^t 1_{\{s < S_M\}} \frac{\mathrm{d}(-M_s)}{M_{s-}}, \ t \geqslant 0.$$

不难验证如果 M 恒正, 那么其 Stieltjes 对数 Log$M$ 也是加泛函. 一般地, $\log M$ 是 $1_{[0,S_M)}$-加泛函. 如果 M 是连续的, 那么两种对数是一样的. 后面我们会看到为什么 Stieltjes 对数是更好的选择.

设 M 是乘泛函, A 是加泛函. 定义

$$M * A := \int_0^\cdot M_s \mathrm{d}A_s;$$

$$M_- * A := \int_0^\cdot M_{s-} \mathrm{d}A_s,$$

那么它们都是 M-加泛函. 让我们验证 $M * A$.

$$\begin{aligned} M * A_{s+t} &= M * A_s + \int_s^{s+t} M_u \mathrm{d}A_u \\ &= M * A_s + \int_0^t M_{s+u} \mathrm{d}A_{s+u} \\ &= M * A_s + M_s \int_0^t M_u \circ \theta_s \mathrm{d}A_u \circ \theta_s \\ &= M * A_s + M_s \cdot (M * A)_t \circ \theta_s. \end{aligned}$$

如果 A 连续, 那么两者没有区别. 显然 $M_- * \log M = 1 - M$. 对任何 M-加泛函 A, 定义其位势算子

$$U_A^\alpha f(x) := \mathbb{E}^x \int_0^\infty \mathrm{e}^{-\alpha t} f(X_t) \mathrm{d}A_t, \ f \in \mathscr{E}_+^*, \ x \in E.$$

当 $A_t = t$ 时, 这就是通常的豫解算子. 不难验证 $U_A^\alpha f$ 是 M-子过程的 α-过分函数.

引理 4.6.2 设 A 是加泛函, M 是恒正的乘泛函, 那么有恒等式

$$U_A^\alpha = U_{M_- * A}^\alpha + U_{\log M}^\alpha U_{M_- * A}^\alpha.$$

证明 取 E 非负可测函数 f, $x \in E$, $\alpha > 0$, 使得 $U_A^\alpha f(x) < \infty$, 那么

$$U_{\log M}^\alpha U_{M_- * A}^\alpha f(x) = \mathbb{P}^x \int_0^\infty \mathrm{e}^{-\alpha t} \frac{\mathrm{d}(-M_t)}{M_{t-}} P^{X_t} \int_0^\infty \mathrm{e}^{-\alpha s} f(X_s) M_{s-} \mathrm{d}A_s$$

$$
\begin{aligned}
&= \mathbb{P}^x \int_0^\infty \frac{\mathrm{d}(-M_t)}{M_{t-}M_t} \int_t^\infty \mathrm{e}^{-\alpha s} f(X_s) M_{s-}\, \mathrm{d}A_s \\
&= \mathbb{P}^x \int_0^\infty \mathrm{d}\frac{1}{M_t} \int_t^\infty \mathrm{e}^{-\alpha s} f(X_s) M_{s-}\, \mathrm{d}A_s \\
&= -\mathbb{P}^x \int_0^\infty \mathrm{e}^{-\alpha s} f(X_s) M_{s-}\, \mathrm{d}A_s \\
&\quad + \mathbb{P}^x \int_0^\infty \frac{1}{M_{t-}} \mathrm{e}^{-\alpha t} f(X_t) M_{t-}\, \mathrm{d}A_s \\
&= U_A^\alpha f(x) - U_{M_- * A}^\alpha f(x).
\end{aligned}
$$

第三个等号是因为 $\mathrm{d}\dfrac{1}{M_t} = \dfrac{\mathrm{d}(-M_t)}{M_{t-}M_t}$. □

当 $A_t = t$ 时, $U_{M_- * A}^\alpha = V^\alpha$, 我们得到下面恒等式: 当 M 恒正时,

$$
U^\alpha = V^\alpha + U_{\log M}^\alpha V^\alpha,
$$

其中 (V^α) 是 M-子过程的豫解. 比较在一般情况下的恒等式

$$
U^\alpha = V^\alpha + P_M^\alpha U^\alpha.
$$

现在我们设 A 是连续加泛函. 定义

$$
R_A := \inf\{t \geqslant 0 : \ A_t > 0\}.
$$

当然 R_A 是停时而且是终止时. A 的精细支撑 $\operatorname{supp}(A)$ 定义为 R_A 的正则点, 即 $\operatorname{supp}(A) := \{x : \mathbb{P}^x(R_A = 0) = 1\}$, 它是精细闭的. R_A 实际上几乎肯定地是 $\operatorname{supp}(A)$ 的首中时. 对任何可测的 f, 定义

$$
(f * A)_t := \int_0^t f(X_s)\mathrm{d}A_s.
$$

定理 4.6.3 支撑 $\operatorname{supp}(A)$ 是最小的近乎可选的精细闭集 K 满足 $1_K * A = A$.

证明 如果 K 是这样的一个集合, 那么

$$
A_t = \int_0^t 1_K(X_s)\mathrm{d}A_s = \int_0^t 1_{\{s > T_K\}} 1_K(X_s)\mathrm{d}A_s,
$$

因此 $R_A \geqslant T_K$, 其中 T_K 是 K 的首中时. 若 $x \notin K$, 那么因为 K 是精细闭的, 故 $\mathbb{P}^x(R_A > 0) \geqslant \mathbb{P}^x(T_K > 0) = 1$, 因此 $x \notin \operatorname{supp}(A)$, 即 $\operatorname{supp}(A) \subset K$. 现在用 R 表示 R_A, 那么因为 $A_R = 0$, 故对任何 t,

$$A_t = A_t + A_R \circ \theta_t = A_{t+R\circ\theta_t},$$

即 A 在 $(t, t + R\circ\theta_t)$ 上是常数, 因此对几乎所有样本

$$\bigcup_{t\in\mathbf{Q}} (t, t + R\circ\theta_t) = \overline{\{t:\ X_t \in \operatorname{supp}(A)\}}^c.$$

而 $\{t:\ X_t \in \operatorname{supp}(A)\}$ 是右闭的, 故它与它的闭包至多差个可列集. 因此

$$1_{\operatorname{supp}(A)} * A = A. \qquad \square$$

用 F 表示 $\operatorname{supp}(A)$. 定义 A 的右连续逆, 对任何 $t \geqslant 0$,

$$\tau_t(\omega) := \inf\{s:\ A_s(\omega) > t\}.$$

显然 τ_t 是停时, $t \mapsto \tau_t$ 是右连续严格递增的, $A_{\tau_t} = t$, $\tau_0 = R_A$ 定义

$$Y_t := X_{\tau_t},\ \widehat{\mathscr{F}_t} := \mathscr{F}_{\tau_t},\ \widehat{\theta}_t = \theta_{\tau_t}.$$

随机过程 (Y_t) 显然是 $\widehat{\mathscr{F}_t}$ 适应的右连续过程且 $(\widehat{\theta}_t)$ 是 (Y_t) 的推移算子族. 事实上, 对任何 $u, t \geqslant 0$,

$$\begin{aligned}\tau_t + \tau_u\circ\theta_{\tau_t} &= \tau_t + \inf\{s : A_s\circ\theta_{\tau_t} > u\} \\ &= \tau_t + \inf\{s : A_{s+\tau_t} > u + t\} = \tau_{u+t}.\end{aligned}$$

因此

$$Y_t\circ\widehat{\theta}_s = X_{\tau_t}\circ\theta_{\tau_s} = X_{\tau_t\circ\theta_{\tau_s}+\tau_s} = X_{\tau_{t+s}} = Y_{t+s}.$$

过程

$$Y = (\Omega, \mathscr{F}, \widehat{\mathscr{F}_t}, Y_t, \widehat{\theta}_t, \mathbb{P}^x)$$

称为是 X 的由连续加泛函 A 诱导的时间变换过程.

定理 4.6.4 X 的由连续加泛函 A 诱导的时间变换过程 Y 是以 F 为状态空间的右过程.

证明 让我们抛开诸多的细节只验证 Y 是右连续简单 Markov 过程且满足 (HD2). 右连续性是显然的, 需验证简单 Markov 性. 对 $\mathscr{F}$ 上有界可测的 Z, 利用 X 的强 Markov 性推出

$$\mathbb{E}^x(Z\circ\widehat{\theta}_t|\widehat{\mathscr{F}_t}) = \mathbb{E}^x(Z\circ\theta_{\tau_t}|\mathscr{F}_{\tau_t}) = \mathbb{E}^{X_{\tau_t}}Z = \mathbb{E}^{Y_t}Z.$$

最后我们来证明 (HD2). 对任何 $\alpha > 0$, 用 $\widehat{U}^\alpha$ 表示 Y 的位势算子, 对 E 上有界可测的 f,

$$\begin{aligned}\widehat{U}^\alpha f(x) &= \mathbb{E}^x\int_0^\infty \mathrm{e}^{-\alpha t}f(Y_t)\mathrm{d}t\\ &= \mathbb{E}^x\int_0^\infty \mathrm{e}^{-\alpha t}f(X_{\tau_t})\mathrm{d}t\\ &= \mathbb{E}^x\int_0^\infty \mathrm{e}^{-\alpha A_t}f(X_t)\mathrm{d}A_t\\ &= U^\alpha_{M*A}f(x),\end{aligned}$$

其中 $M_t := \mathrm{e}^{-\alpha A_t}$ 是恒正的乘泛函. 由引理 4.6.2 知 $\widehat{U}^\alpha f$ 实际上是两个 α-过分函数的差, 因此 $\widehat{U}^\alpha f(X)$ 是右连续的, 从而 $\widehat{U}^\alpha f(Y)$ 也是右连续的. □

如果 $A_t = \int_0^t 1_F(X_s)\mathrm{d}s$, 那么直观地看 A 只在过程在 F 中时增加, A 诱导的时间变换就是简单地抹去 X 不在 F 中的片段.

习　题

1. 如果把定义 4.6.1 条件 (1) 改为 (1′): 对任何 $x \in E$, $t \geqslant 0$ 有 $M_t \geqslant 0$ 且 $\mathbb{E}^x M_t \leqslant 1$, 证明: M 是一个非负上鞅, 称为上鞅乘泛函.
2. 设 M 是乘泛函, 证明: 当 $t \in (0, s)$ 且递减时, $M_{s-t}\circ\theta_t$ 递减.
3. 设 $H \in b\mathscr{F}_t^{0*}$, 证明: $\mathbb{Q}^x(H1_{\{t<\zeta\}}) = \mathbb{E}^x(HM_t)$.
4. 设 M 是右的乘泛函, 证明: 对任何 $t, s \geqslant 0$, 有
$$\mathbb{Q}^x(\zeta > t+s|\mathscr{F}^{0*}_{t+})1_{\{\zeta>t\}} = \mathbb{E}^{X_t}(M_s)1_{\{\zeta>t\}}.$$
5. 设 X 是 $\mathbf{R}$ 上向右一致平移, $T = T_0$, $0 < \beta < 1$. 定义 $M := 1_{[0,T)} + (1-\beta)1_{[T,\infty)}$, 证明 M 是恰好的乘泛函.

6. 证明: 如果 M 恒正, 那么其 Stieltjes 对数 $\log M$ 也是加泛函. 一般地, $\log M$ 是 $1_{[0,S_M)}$-加泛函.
7. 设 A 是 M-加泛函, 证明: $U_A^\alpha f$ 是 M-子过程的 α-过分函数.
8. 设 A 是加泛函, 证明: R_A 是停时而且是终止时.
9. 设 A 是 $\mathbf{R}$ 的右闭子集(即 A 对递减极限封闭), 证明: A 与其闭包至多差个可列集.
10. 加泛函 A 不负荷可测集 F 是指 $1_F * A = 0$, 证明: 连续加泛函不负荷半极集.
11. 设 Y 是 X 的由连续可加泛函 A 诱导的时间变换过程, 证明:

    (1) 如果 f 是 X 过分的, 那么 $f|_F$ 是 Y 过分的;

    (2) 如果 Y 是暂留的且 $F = E$, 那么一个 Y 的过分函数是 X 的过分函数.

参 考 文 献

[1] Bauer, H., *Probability Theory and Elements of Measure Theory,* Academic Press, 1981

[2] Bertoin, J., *Lévy Processes,* Cambridge University Press, 1996

[3] Billingsley, P., *Probability and Measure, John Wiley & Sons,* 1986

[4] Berg, C., Forst, G., *Potential Theory on Locally Compact Abelian Group* , Springer-Verlag, 1973

[5] Blumenthal, R. M., *Getoor, R. K., Markov Processes and Potential Theory,* Academic Press, 1968

[6] Chung, K.L., *Lectures from Markov Processes to Brownian Motion,* Springer-Verlag, New York Heidelberg Berlin, 1982

[7] Chung, K.L., A Course in Probability Theory, Academic Press, 1974

[8] Chung, K.L.,Williams, R.J.,*Introduction to Stochastic Integration,* Birkhauser Boston, Inc., 1983

[9] Dellacherie, C., *Capacités et processus stochastiques,* Springer-Verlag, Berlin Heidelberg New York, 1972

[10] Dellacherie, C., Meyer, P. A., *Probabilities and Potential Vol.* 1,2,3, North-Holland, 1982

[11] Doob, J.L., *Stochastic Processes,* John Wiley & Sons, 1953

[12] Durrett, R., *Brownian motion and Martingale in Analysis,* Wadsworth Inc., 1985

[13] Dynkin, E.B., *Markov Processes,* Springer, Berlin Heidelberg New York, 1965

[14] Feller, W., *An Introduction to Probability Theory and its Applications,* Vol. 1(1959), 2(1970), Wiley & Son

[15] Getoor, R.K., Markov Processes: *Ray Processes and Right Processes, Lecture*

Notes in Math 440, Springer-Verlag, 1975

[16] Getoor, R.K., *Excessive Measures,* Birkhauser, 1990

[17] Halmos, P.R., *Measure Theory,* Springer-Verlag, 1974

[18] Ikeda, N., Watanabe, S., *Stochastic Differential Equations and Diffusion Processes,* North-Holland, 1981

[19] Itô, K., *Lectures on Stochastic Processes,* Tata Institute, Bombay 1961

[20] Itô, K., Mckean Jr., H. P., *Diffusion Processes and Their Sample paths,* Springer-Verlag, 1965

[21] Kallengberg, O., *Foundation of Modern Probability,* Springer-Verlag, 2001

[22] Karatzas, I., Shreve, S. E., *Brownian Motion and Stochastic Calculus,* Springer, 1991

[23] Karlin, S., Taylor, H. M., *A First Course in Stochastic Processes,* Academic Press, 1975

[24] Krylov, N.V., *Introduction to the Theory of Random Processes, Graduate Studies in Math. 43, AMS,* 2002

[25] Laha, R.G., Rohatgi, V.K., *Probability Theory,* John Wiley & Sons, 1979

[26] von Mises, R., *Probability, Statistics and Truth,* Dover Publications, Inc., 1957

[27] Oksendal, B., *Stochastic Differential Equations,* Springer-Verlag, 1995

[28] Parthasarathy, K.R., *Probability Measures on Metric Spaces,* Academic Press, 1967

[29] Revuz, D., Yor, M., *Continuous Martingales and Brownian Motion,* Springer, 1991

[30] Sharpe, M.J., *General Theory of Markov Processes,* Academic Press, Inc., 1990

[31] Varadhan, S. R. S., *Probability Theory, Courant Lecture Notes 7, AMS,* 2001

[32] Yosida, K., *Functional Analysis,* Springer-Verlag, 1980

[33] 王梓坤, 随机过程通论, 1,2, 北京师范大学出版社, 1996

[34] 汪嘉冈, 现代概率论基础, 复旦大学出版社, 1988

[35] 李漳南, 吴荣, 随机过程教程, 高等教育出版社, 1987

[36] 严加安, 测度论讲义, 科学出版社, 1998

[37] 何声武, 汪嘉冈, 严加安, 半鞅与随机分析, 科学出版社, 1995

索　引

随机过程基础：习题与解答

应坚刚

2007 年 12 月 28 日

第一章：概率论基础

§1.1 习　题

1. 设 $\mathscr{A}$ 是原子的集合，那么 $|\mathscr{F}| = 2^{|\mathscr{A}|}$. 原子就是不含有可测真子集的非空可测集.
2. 举例：　$\mathscr{F}_n = \sigma(\{1\}, \cdots, \{n\})$, 那么 $\bigcup_n \mathscr{F}_n$ 可数，不可能是 σ- 代数.
3. $\mathscr{B}(\mathbf{R}) = \sigma(\{(-\infty, x) : x \in Q\})$, 因此它最多与实数等势，而 Lebesgue 可测集全体是与实数的幂集等势的.
5. 取一列不交的 $A_n \in \mathscr{F}_0$, 且 $\bigcup A_n \in \mathscr{F}_0$, 令 $C_n = \bigcup_n^\infty A_i$, 那么 $C_n \in \mathscr{F}_0$ 且 $\lim C_n = \emptyset$, 因此 $\mu(\bigcup A_i) = \sum_1^n \mu(A_i) + \mu(C_{n+1})$, 取极限.
6. (2) 若有 $A_1 \cup N_1 = A_2 \cup N_2$, 要验证 $\mu(A_1) = \mu(A_2)$. 设 $N_i' \in \mathscr{F}$, $N_i \subset N_i'$, $\mu(N_i') = 0$. 记 $N_0 = N_1 \cup N_2$, $N_0' = N_1' \cup N_2'$, 则 $\mu(A_1) = \mu(A_1 \cup N_1') \le \mu(A_1 \cup N_0') = \mu(A_2 \cup N_0') = \mu(A_2)$, 推出 $\mu(A_2) = \mu(A_1)$.
7. (1) μ^* 的定义是

$$\mu^*(A) = \inf\{\sum \mu(B_n) : B_n \in \mathscr{F}, A \subset \bigcup B_n\}.$$

因此 $\mu^* \le \mu^{**}$. 反之对任何 $\epsilon > 0$,

$$\mu^*(A) + \epsilon \ge \sum \mu(B_n) \ge \mu(\bigcup B_n) \ge \mu^{**}(A).$$

(3) 由于 $\mu(\Omega) < \infty$, 故 $\mu_*(A) = \mu(\Omega) - \mu^*(A^c)$.
$\Leftarrow$: 要证对任何 $E \subset \Omega$, 有 $\mu^*(E) \ge \mu^*(A \cap E) + \mu^*(A^c \cap E)$. 不妨设左边有限，对任何 $\epsilon > 0$, 存在 $B \in \mathscr{F}$, $\mu^*(E) + \epsilon \ge \mu(B)$. 因 $\mu^*(\Omega) < \infty$, 故存在 $A_1, A_2 \in \mathscr{F}$ 使得 $A_1 \subset A \subset A_2$ 且 $\mu(A_2 \setminus A_1) < \epsilon$. 则

$$\begin{aligned}\mu^*(E) + \epsilon &\ge \mu(B) = \mu(B \cap A_2) + \mu(B \setminus A_2)\\ &\ge \mu^*(E \cap A) + \mu(B \setminus A_1) - \epsilon\\ &\ge \mu^*(E \cap A) + \mu^*(E \setminus A) - \epsilon.\end{aligned}$$

$\Rightarrow$: 显然 $\mu^*(A) \ge \mu_*(A)$. 反之，如果 $A \in \mathscr{M}$, 则 $\mu(\Omega) \ge \mu^*(A) + \mu^*(A^c)$. 推出 $\mu_*(A) \ge \mu^*(A)$.
8. 单调类定理：对任何 $E \in \mathscr{F}$, 作集类 $[E] := \{F \in \mathscr{F} : E \cup F, E \setminus F, F \setminus E \in \mathscr{F}\}$. 那么 $[E]$ 是单调类. 如果 $E \in \mathscr{A}$, 那么 $\mathscr{A} \subset [E]$, 推出 $m(\mathscr{A}) \subset [E]$, 这里 $m(\mathscr{A})$ 表示 $\mathscr{A}$ 生成的单调类. 这意味着如果 $F \in \mathscr{F}$, 那么 $\mathscr{A} \subset [F]$, 又推出 $m(\mathscr{A}) \subset [F]$. 也就是说 $m(\mathscr{A})$ 对并与差运算封闭，因此，它是 σ- 代数，故而 $\sigma(\mathscr{A}) \subset m(\mathscr{A}) \subset \mathscr{F}$.

10. 一个无限的 σ- 代数其原子的个数肯定是无限多个.
11. $\Rightarrow$: 如果 $A \supset B$, 那么 $A \setminus B = (A^c \cup B)^c$. 如果 A_n 递增，那么 $\lim A_n = \bigcup (A_n \setminus A_{n-1})$. $\Leftarrow$: 因为 $\omega \in \mathscr{A}$, 故它对补封闭. 再证 $\mathscr{A}$ 对不交并封闭. 设 $A, B \in \mathscr{A}$, 不交. 则 $A \cup B = (A^c \setminus B)^c$.
12. (1) 因为连续函数一定是 Borel 可测的. (2) 只需证明当 Ω 是度量空间时, $\mathscr{A}(\Omega)$ 包含所有开集或者闭集. 设 G 是闭集，则 $x \in G$ 当且仅当 $d(x, G) = 0$. 令 $f(x) := d(x, G)$. 则 $G = \{x : d(x, G) = 0\}$.
15. 设 $x \in [-1, 1]$,

$$\lambda \xi^{-1}([-1, x]) = \lambda(\{\theta \in [0, 2\pi] : \sin\theta \le x\}) = 2 \arcsin x.$$

17. 证明 $\mathsf{P}([0, 1/4]) = 1/4$. 其它类似. 首先由测度限制

$$\begin{aligned} &\mathsf{P}([0,1/2]) + \mathsf{P}([0,1/4] \cup [1/2,3/4]) \\ &\quad = \mathsf{P}([0,1/4]) + \mathsf{P}([0,3/4]) = 2\mathsf{P}([0,1/4]) + \mathsf{P}([1/4,3/4]). \end{aligned}$$

由条件得 $\mathsf{P}([0, 1/4]) = 1/4$.
18. (1) 由 Jordan 分解推出. (2) 显然这样的测度如果存在必唯一. 令 $\mu \vee \nu := 1_D \cdot \nu + 1_{D^c} \cdot \mu$. 显然它满足 (a), 另外对任何 A, $\mu \vee \nu(A) = \nu(D \cap A) + \mu(D^c \cap A) \le k(D \cap A) + k(D^c \cap A) = k(A)$, 因此 (b) 满足.
19. 设 $f \in L^1(\mathbf{R}, \mu)$, 则对任何 $\epsilon > 0$, 存在 N 使得 $\mu(|f 1_{\{|x|>N\}}|) < \epsilon$. 而 $f 1_{\{|x|>N\}}$ 可以由简单函数逼近，所以我们只需证明示性函数可以被连续函数逼近. 对任何测度有限的 Borel 集 B, 与 $\epsilon > 0$, 存在开集 G 与闭集 F, 使得 $F \subset B \subset G$ 且 $\mu(G \setminus F) < \epsilon$. 定义

$$f(x) = \frac{d(x, G^c)}{d(x, G^c) + d(x, F)}$$

那么 f 连续， $f|_G = 0$, $f|_F = 1$ 且 $\mu(|1_B - f|) < \epsilon$.
20. 实际上证明

$$\mathscr{F} \subset \bigcup_{J \subset I} \sigma(\{\xi_i : i \in J\}),$$

其中 J 跑遍所有可列子集. 这只需验证右边是 σ- 代数就够了.
21. (参考严加安的测度论讲义) 显然证明第二句话就够了. Put $\mu = \sum_n \mathsf{P}_n / 2^n$, 那么 μ 是概率测度，且 $\mathsf{P}_n \ll \mu$. 定义 $d(A, B) = \mu(A \Delta B)$, $A, B \in \mathscr{F}$. 那么 $(\mathscr{F}, d)$ 是完备度量空间，因为 $\{1_A : A \in \mathscr{F}\}$ 是 $L^1(\mu)$ 中的闭集. 设 $a > 0$, 令

$$L_j = \{A \in \mathscr{F} : \forall n, m >\ge j, |\mathsf{P}_n(A) - \mathsf{P}_m(A)| < a\},$$

由于函数 $A \mapsto \mathsf{P}_n(A)$ 在 $\mathscr{F}$ 上连续，故 L_j 是闭集．由条件知 $\mathscr{F} = \bigcup_j L_j$．由 Baire 纲定理推出，某 L_j 有内点 A. 即存在 $h > 0$ 使得 $\{B : \mu(B\Delta A) < h\} \subset L_j$. 由绝对连续性，存在 $\delta > 0$ 使得 $\mu(C) < \delta$ 蕴含 $\mathsf{P}_i(C) < a$, $1 \le i \le j$. 那么对 $n \ge j$,

$$\begin{aligned}
\mathsf{P}_n(C) &= \mathsf{P}_n(A \cup C) - \mathsf{P}_n(A) + \mathsf{P}_n(A) - \mathsf{P}_n(A \setminus C) \\
&\le |\mathsf{P}_n(A \cup C) - \mathsf{P}_j(A \cup C)| + |\mathsf{P}_j(A \cup C) - \mathsf{P}_j(A)| + |\mathsf{P}_j(A) - \mathsf{P}_n(A)| \\
&+ |\mathsf{P}_n(A \setminus C) - \mathsf{P}_j(A \setminus C)| + |\mathsf{P}_j(A \setminus C) - \mathsf{P}_j(A)| + |\mathsf{P}_j(A) - \mathsf{P}_n(A)| \\
&\le 6a.
\end{aligned}$$

即推出 $B_k \downarrow \emptyset$ 蕴含 $\sup_n \mathsf{P}_n(B_k) \downarrow 0$. 实际上第一句话与第二句话等价.
(另一种证明：赵敏智提供) Let A_n decrease to empty set. We need to show $\lim_k \lim_n \mathsf{P}_n(A_k) = 0$. Suppose that $\lim_k \lim_n \mathsf{P}_n(A_k) = c > 0$. Then when k large enough,

$$9c/8 > \lim_n \mathsf{P}_n(A_k) \ge c.$$

There exist n_1 and k_1 such that $9c/8 > \mathsf{P}_{n_1}(A_{k_1}) > 7c/8$. Then there exists $k_2 > k_1$ such that $\mathsf{P}_{n_1}(A_{k_2}) < c/8$. There exists $n_2 > n_1$ such that

$$9c/8 > \mathsf{P}_{n_2}(A_{k_2}) > 7c/8.$$

This procedure gives two sequences $\{n_i\}$ and $\{k_i\}$ satisfying

$$9c/8 > \mathsf{P}_{n_i}(A_{k_i}) > 7c/8,\ \mathsf{P}_{n_i}(A_{k_{i+1}}) < c/8.$$

Therefore we may assume that $\{\mathsf{P}_n\}$ and $\{A_k\}$ satisfies

$$9c/8 > \mathsf{P}_n(A_n) > 7c/8,\ \mathsf{P}_n(A_{n+1}) < c/8.$$

Set

$$A := \bigcup_{n \ge 1} (A_{2n-1} \setminus A_{2n}).$$

Then

$$\mathsf{P}_{2n-1}(A) \ge \mathsf{P}_{2n-1}(A_{2n-1} \setminus A_{2n}) \ge 7c/8 - c/8 = 6c/8$$

and

$$\mathsf{P}_{2n}(A) \le \mathsf{P}_{2n}(A_1 \setminus A_{2n}) + \mathsf{P}_{2n}(A_{2n+1}) \le 9c/8 - 7c/8 + c/8 = 3c/8.$$

This contradicts to the existence of limit $\lim_n \mathsf{P}_n(A)$.

§1.2 习　　题

2. 利用 Dynkin 引理证明对任何 $A \in \sigma(\mathscr{A})$, 有 $1_A \in \mathscr{H}$.
3. $\Leftarrow$: 显然.　$\Rightarrow$: 如果 $\xi = 1_A$, 那么 $A \in \sigma(\eta)$, 也就是说存在 Borel 集 B 使得 $A = \eta^{-1}(B)$, 因此 $\xi = 1_A = 1_B(\eta)$.
4. 在定理 1.2.5 中已经证明.
5. (1) 由定理 1.2.5 推出. 也可以利用表示定理. 定义 $\phi(f) := \langle f, \nu \rangle$, $f \in L^2(\lambda)$. 那么 $|\phi(f)| \le \|f\|_{L^2(\lambda)}$, 因此存在 g. (2) $\nu_s(B^c) = 0$. $\nu(B) = \int_B g d\mu + g d\nu = \mu(B) + \nu(B)$, 因此 $\mu(B) = 0$. 对任何 D, $\mu(D) = 0$. $\int f(1-g) d\nu = \int f g d\mu$. 令 $f = 1_D$, 那么 $\int_D (1-g) d\nu = 0$, 因此 $\nu_a(D) = \nu(D \cap A) = 0$. (3) $(1-g)\cdot \nu_a = (1-g)\cdot \nu = g \cdot \mu$.
6. (1) 不妨设 $\Omega = [0,1]$, 令 $f(x) := \mathsf{P}(A \cap [0,x])$, 这是连续函数.　(2) 存在 $A_1 \subset A$ 使得 $0 < \mathsf{P}(A_1) < \mathsf{P}(A)$. 取 A_1 与 $A \setminus A_1$ 测度小者，继续这个过程. 其实存在 A 的递减可测子集列 A_n 满足 $\mathsf{P}(A_n) < 1/n$. (3) 先设 $\mu(A) < +\infty$. A 的可测子集族 $\mathscr{A}$ 称为串，如果 (a) $\mathscr{A} \supset \{A_n\}$, 其中 A_n 如上;　(b) $\mathscr{A}$ 中的任何两个集合 A_1, A_2, 或者 $A_1 \subset A_2$ 或者 $A_2 \subset A_1$ (忽略 μ- 零测集). 由 Zorn 引理，取一个极大的串 $\mathscr{A}$, 我们证明 $\{\mathsf{P}(B) : B \in \mathscr{A}\}$ 取遍 $[0, \mathsf{P}(A)]$ 中的所有值. 令

$$a = \sup\{\mathsf{P}(B) : B \in \mathscr{A}, \mathsf{P}(B) \le x\},$$
$$b = \inf\{\mathsf{P}(B) : B \in \mathscr{A}, \mathsf{P}(B) \le x\}.$$

首先上面上 (下) 确界在 $\mathscr{A}$ 中可以得到. 因为存在 B_n 使得 $\mathsf{P}(B_n) \uparrow a$, 由于 $\mathscr{A}$ 是个串，故 B_n 递增，则 $\bigcup B_n \in \mathscr{A}$ 且 $\mathsf{P}(\bigcup B_n) = a$. 设 $\mathsf{P}(A'') = b, \mathsf{P}(A') = a$, 如果 $a < x < b$, 那么取 $0 < c < b - a = \mathsf{P}(A'' \setminus A')$, 存在 $B' \subset A'' \setminus A'$ 使得 $0 < \mathsf{P}(B') < c$, 那么 $\mathsf{P}(A') < \mathsf{P}(A' \cup B') < \mathsf{P}(A'')$, 所以把 $A' \cup B'$ 放在 $\mathscr{A}$ 中仍然是一个串，与 $\mathscr{A}$ 的极大性矛盾. 最后需要证明如果 $\mu(A) = +\infty$, 那么存在 A 的可测子集 A' 使得 $x < \mu(A') < +\infty$. 事实上，A 的可测子集族 $\mathscr{A}$ 称为一个串，如果 $\mathscr{A}$ 中的集合测度有限且任何两个集合有包含关系 (忽略零测集不计). 串的全体满足 Zorn 引理的条件，因此有个极大串，记为 $\mathscr{A}$. 如果 $\mathscr{A}$ 中有测度大于 x 的集合，那么证完. 不然，$b := \sup\{\mu(B) : B \in \mathscr{A}\} \le x$, 由可列可加性，存在 $B \in \mathscr{A}$ 使得 $\mu(B) = b$. 那么 $\mu(E \setminus B) = +\infty$. 存在 $B' \subset E \setminus B$ 使得 $0 < \mu(B') < +\infty$. 定义 $A' = B \cup B'$, 那么 A' 包含了 $\mathscr{A}$ 中的所有集合，$\mu(A') > b$. 因此 $\mathscr{A} \cup \{A'\}$ 也是一个串，与 $\mathscr{A}$ 的极大性矛盾.

另一种方法 (赵敏智提供): 先证明对任何测度有限的集合 A, 与任何 $\epsilon > 0$, A 可以写为有限个测度不超过 ϵ 的不交可测集合的并. 事实上，定义

$$\mu_\epsilon(B) = \sup\{\mu(B') : B' \subset B, 0 < \mu(B') \le \epsilon\}.$$

由 (2) 推出 $\mu(B) > 0$ 蕴含 $\mu_\epsilon(B) > 0$. 存在 $A_1 \subset A$, 使得 $\mu_\epsilon(A)/2 < \mu(A_1) \le \epsilon$. 如果 $\mu(A \setminus A_1) > \epsilon$. 存在 $A_2 \subset A \setminus A_1$ 使得 $\mu(A \setminus A_1)/2 < \mu(A_2) \le \epsilon$. 继续，得到 $\{A_n\}$, 满足 $A_n \subset A \setminus (\bigcup_{i=1}^{n-1} A_i)$, 且

$$\mu_\epsilon(A \setminus (\bigcup_{i=1}^{n-1} A_i)/2 < \mu(A_n) \le \epsilon.$$

令 $B = \bigcup A_n$. $\mu(B) \le \mu(A)$, 且

$$\mu_\epsilon(A \setminus B) \le \mu_\epsilon(A \setminus (\bigcup_{i=1}^{n-1} A_i) \le 2\mu(A_n) \longrightarrow 0,$$

因此 $\mu(A \setminus B) = 0$. 存在 n 使得 $\mu(\bigcup_{i>n} A_i) \le \epsilon$, 故 A 划分为 $A_1, \cdots, A_n, \bigcup_{i>n} A_i$.

8. (参考严加安的测度论讲义) 令 $\mathscr{H}_+^*$ 是指 $\mathscr{H}$ 中的非负递增列的极限全体. 定义 $I^*(f) = \sup\{I(g) : g \in \mathscr{H}, f \ge g \ge 0\}$. 取 $\mathscr{H}$ 中递增收敛于 f 的非负函数列 f_n, 显然 $I^*(f) \ge \lim_n I(f_n)$, 反过来，对任何 $g \in \mathscr{H}$, $0 \le g \le f$, 有 $f_n \wedge g \uparrow g$, 故 $\lim_n I(f_n) \ge \lim_n I(f_n \wedge g) = I(g)$. 因此 $I^*(f) = \lim_n I(f_n)$. $I^*(f)$ 与 f_n 的选择无关.

9. $\Rightarrow$: 反证法，存在 A_n 使得 $\mu(A_n) < n^{-2}$, $\nu(A_n) \ge \epsilon$. 令 $A = \bigcap_n \bigcup_{k>n} A_k$, 不难验证 $\mu(A) = 0$, 但 $\nu(A) > \epsilon$.

10. 存在不交的 Ω_n 使得 $\bigcup_n \Omega_n = \Omega$ 且 $0 < \mu(\Omega_n) < \infty$. 令 $f = \sum_n \frac{1}{2^n \mu(\Omega_n)} 1_{\Omega_n}$.

附加题：设 Ω 是不可数集，$\mathscr{F}$ 是 Ω 的单点子集生成的 σ- 代数. 证明：对角线 $\{(\omega, \omega) : \omega \in \Omega\} \notin \mathscr{F} \times \mathscr{F}$

定义 $\mathscr{F}$ 上的测度 μ, 如果 A 可列，$\mu(A) = 0$, 如果 A^c 可列，$\mu(A) = 1$. 那么 $(\Omega, \mathscr{F}, \mu)$ 是概率空间. 反证法，如果对角线 D 可测，那么 Fubini 定理可用. 这样容易计算 1_D 的累次积分是零. 再算 $\mu \times \mu(D)$, 覆盖 D 的任何矩形可测集列中必然有一个矩形是两个不可列集的乘积，因此 $\mu \times \mu(D) = 1$. 矛盾.

§1.3 习　　题

2. 因为 F 连续，$F^{-1}(x) = \inf\{y : F(y) \le x\} = \inf\{y : F(y) = x\}$, 故 $F(F^{-1}(x)) = x$, 因此 $\mathsf{P}(F(\xi) \le x) = \mathsf{P}(\xi \le F^{-1}(x)) = x$.

4. F 满足 $(\hat{F}(x/\sqrt{n}))^n = \hat{F}(x)$.
18. (2) 设 $a_n > 0$, $\sum_n a_n = 1$, 证明： $\{\sum_{n\in A} a_n : A \subset \mathbf{N}\}$ 是闭集. In fact, put $\mu(\{n\}) = 2^{-n}$ and $d(A,B) = \mu(A\Delta B)$ for $A, B \subset \mathbf{N}$, $g(A) = \sum_{n\in A} a_n$. Then d is a metric on $2^{\mathbf{N}}$. Two things need to check: (1) $(2^{\mathbf{N}}, d)$ is compact; 用对角线法证明：对任何序列，取第一分量都是 0(或 1, 总是可以) 的子列，再其中取第二个分量都是 0(或 1) 的子列，继续，然后拿出对角线列，其 n 个分量从第 n 个元素后都是一样的，所以它收敛. (2) g is continuous. We know that the image of compact set under continuous mapping is compact. That completes the proof. 一般地， $\Omega = A \cup A^c$, P 在 A 上离散，在 A^c 上非原子. 那么 P 的值域等于 $\{\mathsf{P}(B) : B \subset A\} + \{\mathsf{P}(B) : B \subset A^c\}$ 是闭集的和.
25. 先固定 ω_2 对 ω_1 取期望， $\mathsf{E}\xi\eta - \xi(\omega_2)\mathsf{E}\eta - \eta(\omega_2)\mathsf{E}\xi + \xi(\omega_2)\eta(\omega_2) \geq 0$, 然后再对 ω_2 取期望，得 $\mathsf{E}\xi\eta - \mathsf{E}\xi\mathsf{E}\eta \geq 0$.
26. $\xi = \sum_n 1_{A_n}$, $\mathsf{E}\xi = \sum_n \mathsf{P}(A_n)$, 由 Borel-Cantelli 引理， $\mathsf{E}\xi = +\infty$ 蕴含着 $\mathsf{P}(\limsup A_n) = 1$, 即 $\mathsf{P}(\xi = +\infty) = 1$.
27. $\mathscr{F} = \sigma\left(\bigcup_n \sigma(\xi_1, \cdots, \xi_n)\right)$.
28. (赵敏智提供) 设 $\mathscr{A}$ 是尾 σ- 代数. 令 $\xi_n = 1_{A_n}$,

$$\tilde{\mathscr{A}} := \{\bigcup_{n\geq 1}\bigcap_{k\geq n}\{\xi_k = a_k\} : a_k \in \{0,1\}\}.$$

那么 (1) 由条件推出 $\tilde{\mathscr{A}}$ 中所有集合概率为零； (2) 所有集合互不相交，其并为 Ω, 因此 $\tilde{\mathscr{A}}$ 不可列； (3) $\tilde{\mathscr{A}} \subset \mathscr{A}$ 且 $\tilde{\mathscr{A}}$ 中的集合是 $\mathscr{A}$ 中的原子.
It is easy to verify that if $A \in \sigma(X_1, \cdots, X_n, \cdots)$, then $X_n(\omega) = X_n(\omega')$ for all n and $\omega \in A$ implies that $\omega' \in A$.
定义 $\omega \sim \omega'$, 如果存在 k 使得 $n \geq k$ 时 $\xi_n(\omega) = \xi_n(\omega')$. 如果 $A \in \mathscr{A}$, $\omega \sim \omega'$, 那么 $\omega \in A$ implies $\omega' \in A$. (3) follows.
设 $\mathscr{A}$ 是可列生成的， $\mathscr{A} = \sigma(B_1, \cdots, B_n, \cdots)$. Let $B = (\bigcup B_n)^c$. Then it is easy to check that B is an atom in $\mathscr{A}$. By Kolmogorov 0-1 law, we may assume that $\mathsf{P}(B_n) = 0$. Then there exists infinite many elements in $\tilde{\mathscr{A}}$ which are contained in B. This contradicts with the fact that B is an atom in $\mathscr{A}$.

§1.4 习　题

1. 由 $\mathsf{E}\sum|\xi_n| < \infty$, 因此 $\sum\xi_n$ 绝对收敛. 再利用控制收敛定理 $\mathsf{E}\xi = \sum\mathsf{E}\xi_n$.
2. $\mathsf{E}|\xi| = \sum_n \mathsf{E}(|\xi|; n \leq |\xi < n+1)$, $n\mathsf{P}(n \leq |\xi| < n+1) \leq \mathsf{E}(|\xi|; n \leq |\xi| < n+1) < (n+1)\mathsf{P}(n \leq |\xi| < n+1)$, 但是 $\sum_1^\infty n\mathsf{P}(n \leq |\xi| < n+1) = \sum_1^\infty\sum_{k=1}^n \mathsf{P}(n \leq |\xi| <$

$n+1) = \sum_{k=1}^{\infty} \sum_{n\geq k} \cdots = \sum_{k=1}^{\infty} \mathsf{P}(|\xi| \geq k)$.
3. 令 $\xi = \sum 1_{A_n}$, 那么 $G = \{\xi \geq m\}$, 然后 Chebyshev 不等式.
4. (1) 先证明一个命题：如果 $\{\xi_n\}$ 被可积的 η 控制且 $\xi_n \xrightarrow{\text{p}} \xi$, 那么 $\xi_n \xrightarrow{L^1} \xi$. 实际上,

$$\begin{aligned} \mathsf{E}|\xi_n - \xi| &\leq \epsilon + \mathsf{E}(|\xi_n - \xi|; |\xi_n - \xi| > \epsilon) \\ &\leq \epsilon + 2\mathsf{E}(\eta; |\xi_n - \xi| > \epsilon), \end{aligned}$$

然后应用 η 的绝对连续性.
(2) 显然 $(\xi - \xi_n)^+ \leq \xi$ 且 $(\xi - \xi_n)^+ \xrightarrow{\text{p}} 0$, 因此 $(\xi - \xi_n)^+ \xrightarrow{L^1} 0$.
(3) 由 (2) 和条件， $\mathsf{E}|\xi - \xi_n|^- = \mathsf{E}[(\xi - \xi_n)^+ - (\xi - \xi_n)] = \mathsf{E}(\xi - \xi_n)^+ - (\mathsf{E}\xi - \mathsf{E}\xi_n)$ 趋向于 0.
5. 设 F_n 在 D 上收敛，记 $\bar{F}(x) := \lim F_n(x)$, $x \in D$. 用 F 表示 $\bar{F}$ 的右连续化，即

$$F(x) := \lim_{y \in D, y \downarrow\downarrow x} \bar{F}(y), \ x \in \mathbf{R}.$$

现在我们证明 F_n 弱收敛于 F. 给定 $x \in \mathbf{R}$, 取 $x_1 < x < x_2$, $x_1, x_2 \in D$, 显然 $\bar{F}(x_1) \leq F(x) \leq \bar{F}(x_2)$, 那么 $F_n(x_1) \leq F_n(x) \leq F_n(x_2)$, 故

$$\bar{F}(x_1) = \lim F_n(x_1) \leq \liminf F_n(x) \leq \limsup F_n(x) \leq \lim F_n(x_2) = \bar{F}(x_2).$$

如果 F 在 x 点连续, 对任何 $\epsilon > 0$, 一定存在 $x_1' < x < x_2'$ 使得 $F(x_2') - F(x_1') < \epsilon$. 因为 D 稠密，可以做到 $x_1' < x_1 < x_2 < x_2'$, 因此

$$\bar{F}(x_2) - \bar{F}(x_1) \leq F(x_2') - F(x_1') < \epsilon.$$

由此推出 $\lim_n F_n(x)$ 存在等于 $F(x)$.
6. 设 $F_n \xrightarrow{\text{w}} F$. 对任何自然数 N, 令

$$\begin{aligned} F_n^{(N)} &:= 1_{[N,\infty)} + F_n \cdot 1_{[-N,N)}, \\ F^{(N)} &:= 1_{[N,\infty)} + F \cdot 1_{[-N,N)}. \end{aligned}$$

那么 $F_n^{(N)} \xrightarrow{\text{w}} F^{(N)}$. 因两边都是分布函数，对 $f \in C_\infty(\mathbf{R})$, 定理 2.4.9 推出

$$\int f \mathrm{d}F_n^{(N)} \longrightarrow \int f \mathrm{d}F^{(N)}.$$

给定 $\epsilon > 0$, 存在 N, 使得当 $|x| > N$ 时， $|f(x)| < \epsilon$. 而

$$\left| \int f \mathrm{d}F_n^{(N)} - \int f \mathrm{d}F_n \right| \leq f(-N)F_n(-N) + f(N)(1 - F_n(N)) + \int_{|x| \geq N} |f| \mathrm{d}F_n \leq 2\epsilon.$$

因此
$$\left|\int f\mathrm{d}F_n - \int f\mathrm{d}F\right| \le 4\epsilon + \left|\int f\mathrm{d}F_n^{(N)} - \int f\mathrm{d}F^{(N)}\right|.$$

7. 首先，连续的分布函数必定一致连续．对任何 $\epsilon > 0$, 存在 $N > 0$ 和 $\delta > 0$ 使得 $F(-N) < \epsilon$, $1 - F(N) < \epsilon$ 且只要 $x < y$ 是 $[-N, N]$ 中距离小于 δ 的两点，则 $F(y) - F(x) < \epsilon$. 现在取 $-N = x_0 < x_1 < \cdots < x_{u-1} < x_u = N$ 使得连续两点的距离小于 δ. 然后 n 充分大 (只与 ϵ 有关), 使得对所有 i, $F_n(x_i) - F(x_i)| < \epsilon$. 现在可以证明对所有 x, $|F_n(x) - F(x)| < 5\epsilon$. 当 $x < -N$ 时 ($x > N$ 类似),

$$|F_n(x) - F(x)| \le F_n(x) + F_n(x) \le F_n(-N) + F(-N) \le \epsilon + 2F(-N) < 3\epsilon,$$

当 $x \in [-N, N]$, 那么存在 i 使得 $x_i \le x < x_{i+1}$,

$$\begin{aligned}|F_n(x) - F(x)| &\le F_n(x) - F_n(x_i) + |F_n(x_i) - F(x_i)| + F(x) - F(x_i)\\ &\le 2\epsilon + F_n(x_{i+1}) - F_n(x_i)\\ &\le 2\epsilon + |F_n(x_{i+1}) - F(x_{i+1})| + F(x_{i+1}) - F(x_i) + |F(x_i) - F_n(x_i)|\\ &\le 5\epsilon,\end{aligned}$$

8. 由 Hölder 不等式， $\mathsf{E}(|xi_n|; |\xi_n| > a) \le (\mathsf{E}|\xi_n|^p)^{1/p}[\mathsf{P}(|\xi_n| > a)]^{1/q}$. 再看 $\mathsf{P}(|\xi_n| > a) \le \mathsf{E}|\xi_n|^p/a^p$.

12. (参考汪嘉岗的现代概率论基础) (1) 容易验证 $\{S_n/n\}$ 一致可积，故只需证 $(S_n - \mathsf{E}S_n)/n \xrightarrow{\mathrm{p}} 0$. (2) 由一致可积性

$$\mathsf{E}\left|\frac{1}{n}\sum_1^n \xi_i - \xi 1_{\{|\xi_i|<n\}}\right| \le \frac{1}{n}\sum_1^n (|\xi_i|; |\xi_i| \ge n) \to 0,$$

只需证 $\frac{1}{n}\sum_1^n(\xi_i 1_{\{|\xi_i|<n\}} - \mathsf{E}\xi_i 1_{\{|\xi_i|<n\}})$ 依概率收敛于 0. (3) 由 Chebyshev 不等式与独立性

$$\mathsf{P}\left(\frac{1}{n}\left|\sum_1^n(\xi_i 1_{\{|\xi_i|<n\}} - \mathsf{E}\xi_i 1_{\{|\xi_i|<n\}})\right| > \delta\right)$$
$$\frac{1}{\delta^2 n^2}\sum \mathsf{E}(\xi_i 1_{\{|\xi_i|<n\}} - \mathsf{E}\xi_i 1_{\{|\xi_i|<n\}})^2.$$

只需证 $\frac{1}{n^2}\sum_1^n(\xi_i 1_{\{|\xi_i|<n\}})^2$ 极限为 0. (4) 显然 $\frac{1}{n^2}\sum_1^n(\xi_i 1_{\{|\xi_i|<n\}})^2 \le \frac{1}{n}\sup_i \mathsf{E}(\xi_i^2; |\xi_i| < n)$, 而由 Fubini 定理

$$\mathsf{E}(\xi_i^2; |\xi_i| < n) = 2\mathsf{E}(\int_0^\infty 1_{\{|\xi_i|>y\}} y dy; |\xi_i| < n)$$

$$
\begin{aligned}
&= 2\mathsf{E}\int_0^n 1_{\{y<|\xi_i|<n\}} y dy \\
&\le 2\int_0^n y\mathsf{P}(|\xi_i|>y)dy \le 2\int_0^n \mathsf{E}(|\xi_i|;|\xi_i|>y)dy,
\end{aligned}
$$

再由一致可积性推出所需结论.

13. 设 ξ_n 服从 $b(n,x)$. 那么 ξ_n/n 依概率收敛于 x 且 $B_nf(x)=\mathsf{E}f(\xi_n/n)$. 由 f 的一致连续性, $\mathsf{E}|f(\xi_n/n)-f(x)|\le 2\|f\|\mathsf{P}(|\xi_n/n-x|>\delta)+\epsilon$, $\mathsf{P}(|\xi_n-x|>\delta)\le \frac{nx(1-x)}{\delta^2 n^2}\le 1/(4n\delta^2)$.

14. 先证明：若 $(\Omega,\mathscr{F},\mu)$ 是有限测度空间, f 非负可测，那么

$$
\|f\|_{L^n(\mu)} \longrightarrow \|f\|_{L^\infty(\mu)}.
$$

$\le$ 是显然的. 反之，设 $\|f\|_{L^\infty}<\infty$, 那么对任何 $\delta>0$, 存在 $A\in\mathscr{F}$, $\mu(A)>0$ 使得在 A 上 $f\ge\|f\|_{L^\infty}-\delta$. 那么

$$
\|f\|_{L^n}\ge \left(\int_A f^n d\mu\right)^{\frac{1}{n}} \ge (\|f\|_{L^\infty}-\delta)(\mu(A))^{\frac{1}{n}},
$$

因此结论成立. 利用这个结论

$$
(\mathsf{P}(X_n\in A))^{\frac{1}{n}} = \left(\sqrt{\frac{n}{2\pi}}\int_A e^{-\frac{ny^2}{2}}dy\right)^{\frac{1}{n}} \longrightarrow \left\| e^{-\frac{y^2}{2}}\right\|_{L^\infty(1_A dy)}.
$$

推出要证结果.

§1.5 习　　题

1. (1) 设存在 $t_0\ne 0$, 使得 $|\phi(t_0)|=1$. 先设 $\phi(t_0)=1$, 那么 $\mathsf{E}\cos t_0X=1$, 任取 $h\in(0,1)$,

$$
\mathsf{E}\cos t_0X\le h\mathsf{P}(\cos t_0X\le h)+\mathsf{P}(\cos t_0X>h),
$$

推出 $\mathsf{P}(\cos t_0X\le h)=0$, 因此 $\mathsf{P}(\cos t_0X<1)=0$ 或者 $\mathsf{P}(\cos t_0X=1)=1$, 即 X 分布在 $\{x:\cos t_0x=1\}$ 上，格分布的. 一般地，存在 θ_0 使得 $\mathrm{e}^{i\theta_0}\phi(t_0)=1$, 即 $\mathsf{E}\mathrm{e}^{i(\theta_0/t_0+X)t_0}=1$, 因此 $X+\theta_0/t_0$ 是格分布的， X 也是格分布的.

(2) 现在由 (1), $|\phi_X(x)|=1$ 蕴含着存在 θ 使得 $xX+\theta$ 分布在 $2\pi\mathbf{Z}$; $|\phi_X(x')|=1$ 蕴含着存在 θ' 使得 $x'X+\theta'$ 分布在 $2\pi\mathbf{Z}$. 反证法，如果 X 不是常数，那么存在 $k_1,k_2\in\mathbf{Z}$ 使得 X 以正概率分布在 $(2k_1\pi-\theta)/x$ 与 $(2k_2\pi-\theta)/x$ 上. 自然地存在 $k_1',k_2'\in\mathbf{Z}$ 使得

$$
\frac{2k_1\pi-\theta}{x}=\frac{2k_1'\pi-\theta'}{x'},\ \frac{2k_2\pi-\theta}{x}=\frac{2k_2'\pi-\theta'}{x'}.
$$

因此 $(k_1-k_2)/x=(k_1'-k_2')/x'$, 与不可公度的条件矛盾.
2. (1) 设 ϕ 是密度 f 得特征函数. 先设 $f\in C_c^\infty(\mathbf{R})$, 这时分部积分,

$$\phi(x)=\int \mathrm{e}^{ixy}f(y)\mathrm{d}y=\frac{1}{ix}\int e^{ixy}f'(y)\mathrm{d}y,$$

因此 $|\phi(x)|\le \frac{1}{|x|}\|f'\|_{L^1}\longrightarrow 0$. 对一般 f, 和任何 $\epsilon>0$, 存在 $g\in C_c^\infty(\mathbf{R})$ 使得 $\|f-g\|_{L^1}<\epsilon$. 然后

$$|\phi(x)|\le\|f-g\|+|\int \mathrm{e}^{ixy}g(y)\mathrm{d}y|$$

极限是零. (2) 应用 Fubini 定理

$$\begin{aligned}\int\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2\sigma^2}}|\phi(t)|^2\mathrm{d}t&=\int_{\mathbf{R}^2}f(x)f(y)\mathrm{d}x\mathrm{d}y\int_{\mathbf{R}}\int\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{t^2}{2\sigma^2}}\mathrm{e}^{it(x-y)}\mathrm{d}t\\&=\int_{\mathbf{R}^2}f(x)f(y)\sigma\mathrm{e}^{-\frac{\sigma^2}{2}(x-y)^2}\mathrm{d}x\mathrm{d}y\\&=\int_{\mathbf{R}^2}\mathrm{e}^{-z^2/2}f(z/\sigma+y)f(y)\mathrm{d}z\mathrm{d}y,\end{aligned}$$

先设 f 还是连续且有界的, 当 $\sigma\longrightarrow\infty$ 时, 右边的极限是 $\sqrt{2\pi}\|f\|_{L^2}$, 由 Fatou 引理推出 $|\phi|$ 是平方可积的. 再用控制收敛定理推出结论. 最后用逼近的方法证明结论对一般的密度函数成立.
13. 设 ϕ 是 ξ 的特征函数, 如果 ϕ 在 0 点二次可导, 则 ξ 平方可积且 $\mathsf{E}\xi^2=-\phi''(0)$. 事实上, 由 Fatou 引理

$$\begin{aligned}-\phi''(0)&=\lim_{h\to 0}\frac{2\phi(0)-\phi(h)-\phi(-h)}{h^2}\\&=\lim_{h\to 0}\frac{1}{h^2}\int 2(1-\cos xh)\mathrm{d}F(x)\\&\ge 2\int\liminf\frac{1-\cos xh}{h^2}\mathrm{d}F(x)=\int x^2\mathrm{d}F(x)=\mathsf{E}\xi^2,\end{aligned}$$

因此 ξ 平方可积.
如果 $\phi(x)=\mathrm{e}^{-|x|^n}$ 是特征函数, 那么 $\phi''(0)=0$, 因此得 $\mathsf{E}\xi^2=0$, 即 $\xi=0$, 矛盾.
14. 因为 $(X+Y)+(X-Y)=2X$, 故 $\phi(2x)=\phi(x)^2\cdot\phi(x)\cdot\phi(-x)=\phi(x)^2\cdot|\phi(x)|^2$. (1) ϕ 不会等于 0. 如果 $\phi(a)=0$, 那么推出 $\phi(a/2)=0$, 故 $\phi(a/2^n)=0$ 推出 $\phi(0)=0$ 矛盾. (2) 令 $p(x)=\phi(-x)/\phi(x)$, 那么 $p(2x)=p(x)^2$. 故 $p(x)=(p(x/2^n))^{2^n}$. 而 $\phi(x)=1-\frac{1}{2}x^2+o(x^2)$, 那么 $p(x)=1+o(x^2)$. 因此 $p(x)=\lim_n(p(x/2^n))^{2^n}=1$, 即 ϕ 是实的. (3) 现在 $\phi(x)=\phi(x/2)^4=\phi(x/2^n)^{4^n}\longrightarrow\exp\{-\frac{1}{2}x^2\}$.

16.

$$\int_{-T}^{T}\frac{\mathrm{e}^{-iax}-\mathrm{e}^{-ibx}}{ix}\mathsf{E}\mathrm{e}^{i\xi x}\mathrm{d}x=\mathsf{E}\int_{-T}^{T}\frac{\mathrm{e}^{i(\xi-a)x}-\mathrm{e}^{i(\xi-b)x}}{ix}\mathrm{d}x$$
$$=\mathsf{E}\int_{-T}^{T}\frac{\sin x(\xi-a)-\sin x(\xi-b)}{x}\mathrm{d}x$$

利用恒等式

$$\lim_{T\to\infty}\int_{-T}^{T}\frac{\sin ax}{x}\mathrm{d}x=\pi\mathrm{sgn}(a)$$

得

$$\lim_{T}\int_{-T}^{T}\frac{\mathrm{e}^{-iax}-\mathrm{e}^{-ibx}}{ix}\mathsf{E}\mathrm{e}^{i\xi x}\mathrm{d}x=\pi\mathsf{E}[\mathrm{sgn}(\xi-a)-\mathrm{sgn}(\xi-b)]$$
$$=\pi[\mathsf{P}(\xi=b)+2\mathsf{P}(\xi\in(a,b))+\mathsf{P}(\xi=a)].$$

§1.6 习　　题

1. 存在 $[0,1]$ 上 Borel 可测的 g 使得 $\mathsf{E}(X|Y)=g(Y)$, 由定义 $\mathsf{E}(X;Y\le x)=\mathsf{E}(g(Y);Y\le x)$, 即 $\int_{-x}^{x}X(w)dw=\int_{-x}^{x}g(|w|)dw=2\int_0^x g(w)dw$. 求导 $X(x)+X(-x)=2g(x)$, 因此 $\mathsf{E}(X|Y)=\frac{1}{2}(X(|w|)+X(-|w|))$.

2. 如果 $a_1=a_2=0$, 那么 $\mathsf{E}(X|Y)=\rho\sigma_1 Y/\sigma_2$.

4. $\mathsf{E}(X;Y\in A,Z\in B)=\mathsf{E}(X;Y\in A)\mathsf{P}(Z\in B)=\mathsf{E}(\mathsf{E}(X|Y);Y\in A)\mathsf{P}(Z\in B)=\mathsf{E}(\mathsf{E}(X|Y);Y\in A,Z\in B)$.

6. $\mathsf{E}(\xi_1|S_n,S_{n+1},\cdots)=\mathsf{E}(\xi_1|S_n,\xi_{n+1},\cdots)=\mathsf{E}(\xi_1|S_n)$.

7. $\mathsf{E}(U\le X;X<a)=\int_{u\le x,x<a}p_1(u)p_2(x)dudx=\int_{x<a}\Phi(x)p_2(x)dx=\mathsf{E}(\Phi(X);X<a)$, 因此 $\Phi(X)=\mathsf{P}(U\le X|X)$. $\mathsf{E}\Phi(X)=\mathsf{P}(U\le X)$

10. g 关于 $\mathscr{A}$ 可测,

$$\mathsf{E}(f,A_1\times\Omega_2)=\int_{A_1\times\Omega_2}fd\mathsf{P}_1\times\mathsf{P}_2=\int_{A_1}gd\mathsf{P}_1.$$

12. $\mathsf{E}(|\mathsf{E}(\xi|\mathscr{A})|;|\mathsf{E}(\xi|\mathscr{A})|>N)\le\mathsf{E}(|\xi|;\mathsf{E}(|\xi||\mathscr{A})>N)$, $\mathsf{P}(\mathsf{E}(|\xi||\mathscr{A})>N)\le\mathsf{E}|\xi|/N$.

13. 由 Jensen 不等式 $\mathsf{E}|X+Y|=\mathsf{E}(\mathsf{E}|X+Y||X)\ge\mathsf{E}|\mathsf{E}(X+Y|X)|=\mathsf{E}|X|$.

第二章：随机过程基础

§2.1 习　　题

1. 类似 §1.1 习题 20.
2. 由 1 推出.
6. $\mu_{(t,s)}(f) = \mathsf{E}f(X_s, X_t)$. $\Leftarrow$: 存在 $\mathbf{R}^2$ 上对角线等于零的有界连续函数使得 $\mathsf{P}(|X_t - X_s| > \delta) \le \mathsf{E}f(X_s, X_t) \to \mathsf{E}f(X_t, X_t) = 0$. $\Rightarrow$: 设 $(s,t) \longrightarrow (u,v)$, 则 $X_s \xrightarrow{\mathrm{p}} X_u$, $X_t \xrightarrow{\mathrm{p}} X_v$, 因此 $(X_s, X_t) \xrightarrow{\mathrm{p}} (X_u, X_v)$.
7. 因为依概率收敛的极限与几乎处处收敛的极限是几乎处处相等的.
8. 类似习题 6.

§2.2 习　　题

3. 不对. 设 Ω 是正整数集合，定义 $\mathscr{F}_n = \sigma(\{1\}, \cdots, \{n\})$, 再定义 $\mathscr{F}_n$ 上概率 $\mathsf{P}_n(\{1,\cdots,n\}) = 0$, $\mathsf{P}_n(\{k : k > n\}) = 1$, 满足题中所述条件，但不存在 P 使得 $\mathsf{P}|_{\mathscr{F}_n} = \mathsf{P}_n$. 因为如果存在，由可列可加性， $\mathsf{P}(\Omega) = 0$.
4. 反证法. 假若有，不妨假设它们是有界的，因为两个随机变量独立蕴含着它们在 $L^2[0,1]$ 上正交，而 $L^2[0,1]$ 是可分 Hilbert 空间，矛盾.

§2.3 习　　题

1. 要证明 $\mu(A - x)$ 关于 x 可测. 如果 f 有界连续，那么 $\int f(y-x)\mu(dy)$ 关于 x 连续. 而连续函数生成的 σ- 代数就是 Borel σ- 代数.
4. 设 (3) 为：对 $A \in \mathscr{G}_t$, $\mathsf{P}(A|\mathscr{F}_t) = \mathsf{P}(A|X_t)$. 首先由 Dynkin 定理证明 (M1) 等价于 (3). (1) $\Rightarrow$ (2): 要证 $\mathsf{P}(B \cap A) = \mathsf{E}(\mathsf{P}(B|X_t), A)$. 这由 (1) 两边取期望得到. (2) $\Rightarrow$ (1): (2) 推出 $\mathsf{P}(A \cap B|\mathscr{G}_t) = 1_A \mathsf{P}(B|X_t)$, 两边对 X_t 取条件期望.
5. C-K: $\int P_s(x,y)dyP_t(y,z) = P_{t+s}(x,z)$. 那么

$$\frac{1}{2\pi\sigma_s\sigma_t}\int \exp\left(-\frac{(y-m_s x)^2}{2\sigma_s^2} - \frac{(z-m_t y)^2}{2\sigma_t^2}\right)dy = \frac{1}{\sqrt{2\pi}\sigma_{t+s}}\exp\left(-\frac{(z-m_{t+s}x)^2}{2\sigma_{t+s}^2}\right).$$

作个积分变换：左边等于

$$\frac{1}{m_t}\frac{1}{2\pi\sigma_s\sigma_t/m_t}\int \exp\left(-\frac{y^2}{2\sigma_s^2} - \frac{(z/m_t - m_s x - y)^2}{2(\sigma_t/m_t)^2}\right)dy,$$

由再生性知它等于

$$\frac{1}{m_t}\frac{1}{\sqrt{2\pi(\sigma_s^2 + (\sigma_t/m_t)^2)}}\exp\left(-\frac{(z/m_t - m_s x)^2}{2(\sigma_s^2 + (\sigma_t/m_t)^2)}\right).$$

由此推出两个方程： $m_{s+t}=m_s m_t$, $\sigma_{s+t}^2=m_t^2(\sigma_s^2+(\sigma_t/m_t)^2)$. 存在 a 使得 $m_t=e^{at}$, 第二个方程推出 $\sigma_0=0$, 存在 b 使得 $d\sigma_t^2/dt=bm_t^2$, 因此 $\sigma_t^2=b(e^{2at}-1)/2a$.

§2.4 习　　题

2. $\Leftarrow$: $\tau_y\circ\theta_k+k\ge\tau_y$, 因此 $\{\tau_y\circ\theta_k<\infty\}\subset\{\tau_y<\infty\}$. $\Rightarrow$: $\limsup\{X_n=y\}\subset\{\tau_y\circ\theta_k<\infty\}$.

3. 因为 X 被吸收前与无限制随机游动一样. 可以直接计算一个无限制随机游动有限步到达 0 和 r 的概率. 对任何 $x,y\in\mathbf{Z}$, 令 $\phi_{x,y}$ 是 T_y 在概率空间 $(W,\mathscr{B},\mathsf{P}^x)$ 上的母函数

$$\phi_{x,y}(t):=\mathsf{E}^x t^{T_y}.$$

首先由空间平移不变性

$$\begin{aligned}\phi_{x,y}(t)&=\mathsf{E}^x t^{T_y}=\mathsf{E}^0 t^{T_y\circ\gamma_x}=\mathsf{E}^0 t^{T_{y-x}}=\phi_{0,y-x}(t)\\ \phi_{x,x}(t)&=\phi_{0,0}(t)=\mathsf{E}^0 t^{T_0}=1.\end{aligned}$$

设 $x>0$, 因为随机游动每次移动一个单位，故从 0 出发时通过 x 必须先通过 1, 即 $T_x\ge T_1$, 这时 $T_x=T_x\circ\theta_{T_1}+T_1$, 因此由强 Markov 性,

$$\begin{aligned}\phi_{0,x}(t)&=\mathsf{E}^0 t^{T_x}=\mathsf{E}^0 t^{T_1+T_x\circ\theta_{T_1}}\\ &=\mathsf{E}^0 t^{T_1}\cdot t^{T_x}\circ\theta_{T_1}\\ &=\mathsf{E}^0(t^{T_1}\mathsf{E}^{s_{T_1}}t^{T_x})\\ &=\mathsf{E}^0 t^{T_1}\cdot\mathsf{E}^1 t^{T_x}=\phi_{0,1}(t)\cdot\phi_{1,x}(t).\end{aligned}$$

记 $\phi:=\phi_{0,1}$, 那么 $\phi_{0,x}=\phi^x$. 另一方面，从 0 出发也必有 $T_x\ge 1$, 这时 $T_x=T_x\circ\theta_1+1$, 由 Markov 性

$$\begin{aligned}\phi_{0,x}(t)&=\mathsf{E}^0 t^{T_x}=\mathsf{E}^0 t^{1+T_x\circ\theta_1}\\ &=t\mathsf{E}^0(\mathsf{E}^{s_1}t^{T_x})\\ &=t((\mathsf{E}^1 t^{T_x})p+(\mathsf{E}^{-1}t^{T_x})q)\\ &=tp\phi_{0,x-1}(t)+tq\phi_{0,x+1}(t).\end{aligned}$$

令 $x=1$, 我们得

$$\phi(t)=tp+tq\phi(t)^2,$$

因 $\phi(t) \le 1$, 故解得

$$\phi(t) = \frac{1-\sqrt{1-4t^2pq}}{2tq}.$$

由对偶性得

$$\phi_{0,-1}(t) = \mathsf{E}^0 t^{T_{-1}} = \frac{1-\sqrt{1-4t^2pq}}{2tp}.$$

由此推出， 0 点出发在有限步内到达 1 的概率为

$$\mathsf{P}^0(T_1 < +\infty) = \phi(1) = \frac{1-|p-q|}{2q} = \begin{cases} 1, & p \ge q; \\ \frac{p}{q}, & p < q. \end{cases}$$

即当随机移动偏向右时，它必定在有限时间内到达 1.

4. 简单.

5. 由 4 推出.

6. (1) 因为 $P_n = P_{n-1}P$, 故 $p_{0,0}^{(n)} = p_{0,0}^{(n-1)}p_{0,0} + (1-p_{0,0}^{(n-1)})p_{1,0}$. (2) 令 $q = p_{0,0} - p_{1,0}$, 那么 $p_{0,0}^{(n)} = p_{1,0}(1+q+\cdots+q^{n-1}) + q^n$.

7. (1) 由习题 1 应用反证法. (2) $\Leftarrow$: 可以验证一个状态 x 可以到达的所有状态是个闭集. (3) 常返状态可达的状态肯定是常返的.

10. (1) 一个常返状态可以到达的状态是常返的，并且可以返回. 另外零常返状态是互达等价类. 如果有限状态 Markov 链有零常返状态，那么可以假设它不可分且都是零常返的. 这时可以证明对任何两状态 x, y, $p_{x,y}^{(n)}$ 收敛于零. 这与有限性矛盾. (2) 因为 $\sum_y p_{x,y}^{(n)} = 1$, 故 $\sum_y \sum_n p_{x,y}^{(n)} = +\infty$, 存在 y 使得 $\sum_n p_{x,y}^{(n)} = +\infty$, 由习题 5, y 常返.

11. X_n 的转移函数是 $p_{0,j} = 1_{j=r}$,

$$p_{i,j} = \begin{cases} p, & j = r-i+1; \\ 1-p, & j = r-i; \\ 0, & \text{else}. \end{cases}$$

解方程得平稳分布为

$$\pi = \left(\frac{1-p}{1-p+r}, \frac{1}{1-p+r}, \cdots, \frac{1}{1-p+r}\right).$$

淋湿的概率极限为 $\pi_0 p$.

§2.5 习　题

9. $(S_1,\cdots,S_n)$ 的联合密度为 $\lambda^n e^{-\lambda y_n}1_{\{0<y_1<\cdots<y_n\}}$. S_n 的密度为 $e^{-\lambda t}\frac{(\lambda t)^n}{n!}$ 条件密度为 $n!1_{\{0<y_1<\cdots<y_{n-1}<t\}}/t^n$.
10. $X>t$ 当且仅当 μ 在 $B(t)$ 上无负荷. $\mathsf{P}(X>t)=\mathsf{P}(\mu(B(t))=0)=e^{-\lambda m(B(t))}$, 其中 $B(t)$ 是原点为中心半径为 t 的圆.
11. 只要证明 $\mathsf{E}\sum_{0\le s\le 1}\Delta N_s^1\Delta N_s^2=0$. 取 $0<t_0<\cdots<t_n=1$,

$$\mathsf{E}\sum_k(N_{t_k}^1-N_{t_{k-1}}^1)(N_{t_k}^2-N_{t_{k-1}}^2)=\sum_k\lambda_1\lambda_2(t_k-t_{k-1})^2.$$

取极限.

§2.6 习　　题

6. 与 Kolmogorov 0-1 律类似证明.
7. $\sup_t B_t\ge\limsup_{t\to\infty}B_t$, 对任何 n, 令 $A=\{\limsup B_t>n\}$, 那么 $A\in\bigcap_{t>0}\sigma(\{B_s:s>t\})$, 因此 $\mathsf{P}(A)=0$ 或者 1. 但是

$$\mathsf{P}(A)=\mathsf{E}\limsup 1_{\{B_t>n\}}\ge\limsup\mathsf{P}(B_t>n)=1/2,$$

因此 $\mathsf{P}(A)=1$, 即 $\mathsf{P}(\limsup B_t=+\infty)=1$.
8. $\mathsf{P}(\limsup\{B_{t_k}>0\})\ge\limsup\mathsf{P}(B_{t_k}>0)=1/2$, 由 0-1 律 $\mathsf{P}(\limsup\{B_{t_k}>0\})=1$.
12. $\mathsf{E}X_tX_s=e^{-t-s}\cdot e^{2t\wedge s}=e^{-|t-s|}$.
13. 由定理 2.2.3 推出 $\mathsf{P}^*(C)=1$, 再由定理 2.2.2 说明除空集外 C 不能有其它可测子集，因此 $\mathsf{P}_*(C)=0$.
14. Brown 运动还可以用 Fourier 级数的方法产生. 设 $\{\xi_n:n\ge 0\}$ 是概率空间 $(\Omega,\mathscr{F},\mathsf{P})$ 上的都服从标准正态分布的独立随机变量序列. (说明其存在性.) 函数 $\{1/\sqrt{\pi},\sqrt{2/\pi}\cos x,\cdots,\sqrt{2/\pi}\cos nx,\cdots\}$ 是 $L^2([0,\pi]$ 的标准正交基，依次记为 $\{e_n:n\ge 0\}$. 对任何 $f\in L^2([0,\pi])$, 令

$$H(f):=a_0\xi_0+a_1\xi_1+a_2\xi_2+\cdots,$$

其中 $\{a_n\}$ 是 f 的 Fourier 系数: $a_n:=\langle f,e_n\rangle$. 显然

$$\mathsf{E}H(f)^2=\sum_{n\ge 0}a_n^2=\int f^2(x)\mathrm{d}x,$$

即 H 是 $L^2([0,\pi])$ 到 $L^2(\Omega,\mathscr{F},\mathsf{P})$ 的一个等距嵌入. 对 $t\in[0,\pi]$, 显然 $1_{[0,t]}$ 的 Fourier 级数为

$$1_{[0,t]}(x)=\frac{t}{\sqrt{\pi}}+\sqrt{\frac{2}{\pi}}\sum_{n\ge 1}\frac{\sin nt}{n}\cdot\cos nx.$$

证明：

$$X_t := H(1_{[0,t]}) = \frac{t}{\sqrt{\pi}}\xi_0 + \sqrt{\frac{2}{\pi}}\sum_{n\geq 0}\sum_{m=2^{n-1}}^{2^n-1}\frac{\sin mt}{m}\xi_m.$$

在 $t\in[0,\pi]$ 以概率 1 一致收敛．因此 (X_t) 是连续过程，再证明它在 $[0,\pi]$ 上是 Brown 运动.

解答 (参考 Ito-Mckean 的 Diffusion processes and their sample paths): 令

$$s_{m,n}(t) = \sum_{m}^{n-1}\frac{\sin kt}{k}\xi_k,\ t_{m,n} = \max_{0\leq t\leq\pi}|s_{m,n}(t)|.$$

那么

$$t_{m,n}^2 \leq \max\left|\sum_{m}^{n-1}\frac{e^{ikt}}{k\xi_k}\right|^2 \leq \sum_{m}^{n-1}\frac{\xi_k^2}{k^2} + 2\sum_{l=1}^{n-m-1}\left|\sum_{j=m}^{n-l-1}\frac{\xi_j\xi_{j+l}}{j(j+l)}\right|,$$

然后

$$\begin{aligned}
\mathsf{E}t_{m,n}^2 &\leq \sum_{m}^{n-1}\frac{1}{k^2} + 2\sum_{l=1}^{n-m-1}\left(\mathsf{E}\left|\sum_{j=m}^{n-l-1}\frac{\xi_j\xi_{j+l}}{j(j+l)}\right|^2\right)^{\frac{1}{2}} \\
&\leq \sum_{m}^{n-1}\frac{1}{k^2} + 2\sum_{l=1}^{n-m-1}\left(\sum_{j=m}^{n-l-1}\frac{1}{j^2(j+l)^2}\right)^{\frac{1}{2}} \\
&\leq \frac{n-m}{m^2} + 2(n-m)(\frac{n-m}{m^4})^{\frac{1}{2}}, \\
\mathsf{E}t_{m,n} \leq (\mathsf{E}t_{m,2m}^2)^{\frac{1}{2}} &\leq (\frac{1}{m} + 2/\sqrt{m})^{\frac{1}{2}} \leq 3m^{-\frac{1}{4}}, \\
\mathsf{E}\sum_{n\geq 1}t_{2^{n-1},2^n} &\leq \sum_{n\geq 1}\mathsf{E}t_{2^{n-1},2^n} < \infty.
\end{aligned}$$

第三章：随机分析基础

§3.1 习　　题

3. $\tau = \inf\{s : X_s \neq X_{s-}\}$. $A = \{\tau\geq t\} = \{\tau < t\}^c$.

5. 令 $s=\tau(w)$. 则 $\mathscr{F}_s \subset \{A : w, w'\in A \text{ or } w, w'\notin A\}$. 事实上, 记右边为 $\mathscr{H}$. 容易验证 $\mathscr{H}$ 是 σ- 代数, 取任何 $B\subset E, t<s$, 因 $X_t(w) = X_t(w')$, 故 $\{X_t\in B\}\in\mathscr{H}$.

因此 $\mathscr{F}_s \subset \mathscr{H}$. 然后 $\{\tau \le s\} \in \mathscr{F}_s$, 推出 $w' \in \{\tau \le s\}$, 即 $\tau(w') \le s$, 同理推出 $w' \notin \{\tau < s\}$ 即 $\tau(w') \ge s$.

6. 因为 σ 是 $\mathscr{F}_\tau$ 可测的，故对任何 t, $\{\sigma \le t\} \cap \{\tau \le t\} \in \mathscr{F}_t$. 而因为 $\sigma \ge \tau$, 故左边等于 $\{\sigma \le t\}$.

7. 定义 $\mathscr{F}_{\tau+} = \{A : A \cap \{\tau \le t\} \in \mathscr{F}_{t+}, \forall t\}$. (1) 证明： $\mathscr{F}_{\tau+} = \{A : A \cap \{\tau < t\} \in \mathscr{F}_t, \forall t\}$. $A \in \mathscr{F}_{\tau+}$, $A \cap \{\tau < t\} = \bigcup_n A \cap \{\tau \le t - 1/n\} \in \mathscr{F}_t$. 反之，$A \cap \{\tau \le t\} = \bigcap_n A \cap \{\tau < t + 1/n\}$. (2) 设 $\sigma > \tau$, $A \in \mathscr{F}_{\tau+}$, $A \cap \{\sigma \le t\} = \bigcup_{r \in Q} A \cap \{\sigma \le t\} \cap \{\sigma > r > \tau\} = \bigcup_{r \in Q} A \cap \{\tau < r\} \cap \{r < \sigma \le t\} \in \mathscr{F}_t$. 因此 $A \in \mathscr{F}_\sigma$. (3) 再证明 $\mathscr{F}_{\tau+} = \bigcap_n \mathscr{F}_{\tau+1/n}$.

9. 利用 Poisson 过程的拟左连续性：参考定理 4.3.1 的证明.

概率论
习题与解答

应坚刚

2010 年 1 月 11 日

第一章　初等概率论

§1.1　集合与计数

习　　题

1. 证明 De Morgan 公式与定理 1.1.1.
2. 设 f 是集合 X 到 Y 的映射，$\mathscr{B}$ 是 Y 的一个子集类，定义 X 的子集类

$$f^{-1}(\mathscr{B}) := \{f^{-1}(B):\ B \in \mathscr{B}\}.$$

证明：如果 $\mathscr{B}$ 对补运算 (对应地，并运算，交运算) 封闭，那么 $f^{-1}(\mathscr{B})$ 也对补运算 (对应地，并运算，交运算) 封闭.

3. 设 f 是集合 X 到 Y 的映射，$\mathscr{A}$ 是 X 的一个子集类，定义 Y 的子集类

$$\mathscr{B} := \{B \subset Y : f^{-1}(B) \in \mathscr{A}\}.$$

证明：如果 $\mathscr{A}$ 对补运算 (对应地，并运算，交运算) 封闭，那么 $\mathscr{B}$ 也对补运算 (对应地，并运算，交运算) 封闭.

4. 一个 0-1 序列是指取值是 0 或者 1 的数列. 证明：至多只有有限多元素不等于 0 的 0-1 序列的集合是可列的.
5. 从标准的 52 张扑克牌中取五张牌，求下列情况各有多少不同选择.　(1) 正好是个顺子；　(2) 正好是个同花顺子；　(3) 四种花色都出现.

►(1) 9×4^5; (2) 36; (3) $4 \times \binom{13}{2} \times 13^3$

6. 满足条件 $x_1 + \cdots + x_r = n$ 的 非负整数值向量 $(x_1, \cdots, x_r)$ 有多少个？

►设

$$A = \{(x_1, \cdots, x_r) : x_1 + \cdots + x_r = n, x_1 \geq 0, \cdots, x_r \geq 0\};$$
$$B = \{(x_1, \cdots, x_r) : x_1 + \cdots + x_r = n + r, x_1 > 0, \cdots, x_r > 0\}.$$

令 $\phi(x_1, \cdots, x_r) = (x_1 + 1, \cdots, x_r + 1)$. 那么 ϕ 是 A, B 间的一一对应. 因此 $|A| = |B|$.

7. 设 T_n 是集合 $\{1, 2, \cdots, n\}$ 不同划分的数目，如 $T_1 = 1, T_2 = 2$. 证明

$$T_{n+1} = 1 + \sum_{k=1}^{n} \binom{n}{k} T_k.$$

►应用数学归纳法. 在集合 $\{1,2,\cdots,n+1\}$ 的划分中, 看元素 $n+1$ 所在的组 A, 设 A 中除 $n+1$ 外的元素个数为 k 的划分的全体为 A_k, 那么 $0\le k\le n$. 这 k 个元素的取法有 $\binom{n}{k}$ 中, A 外剩下的 $n-k$ 个元素有 T_{n-k} 种不同划分法. 因此由乘法和加法原理

$$T_{n+1}=\sum_{k=0}^{n}|A_k|=|A_n|+\sum_{0\le k<n}\binom{n}{k}T_{n-k}=1+\sum_{k=1}^{n}\binom{n}{k}T_k.$$

8. 设 A,B 是两个有限集合，那么 $|A\cup B|=|A|+|B|-|A\cap B|$.
9. 求 x_i 取 0 或 1 并满足 $\sum_{i=1}^{n}x_i\ge k$ 的向量 $(x_1,\cdots,x_n)$ 的个数.
10. 求下列情况下 k 对夫妇的坐法有多少种？要求夫妻坐位相邻. (1) 排成一排的 $2n$ 个坐位； (2) 围绕一圆桌的 $2n$ 个坐位.
11. 求从 $1,2,\cdots,n$ 中取 k 个不重合 r- 组的取法种数，其中 r- 组是指 r 个连续整数.
12. 证明：可列个可列集的并仍然是可列的.
13. 证明：代数数集合是可列的.
14. (Schröder-Bernstein, 1898) 如果集合 A, B 互相与对方的一个子集一一对应, 证明： $|A|=|B|$. 因此 $[0,1]$ 与 $(0,1)$ 之间存在一个一一对应. 请写出他们之间的一个一一对应.

§1.2 古典概率模型

习题解答

1. 设有 A,B,C 三个事件，请用集合运算表示下列事件.
   (a) 仅有 A 出现；
   (b) A,B 都出现， C 没有出现；
   (c) 三个事件都出现；
   (d) 至少一个事件出现；
   (e) 至少两个事件出现；
   (f) 仅有一个事件出现；
   (g) 仅有两个事件出现；
   (h) 不多于两个事件出现.
2. n 个球随机地放在 n 个盒子里，求恰有一个空盒的概率.

   ►
   $$\frac{n(n-1)\cdot n!/2!}{n^n}$$

3. 8 部车随机地停在一列 12 个位置的停车场内，求四个空位恰好连在一起的概率.

   ▶
   $$\frac{9}{\binom{12}{4}} = 1/55.$$

4. 在一副经充分洗匀的扑克牌 (这里及后面提到的扑克排都是指不包括 Joker 的 52 张牌) 中，求任何两个 A 都不相邻的概率.

   ▶四张牌分开 48 张牌，分母是可以相邻，分子是不能相邻.
   $$\frac{48! \cdot 49 \cdot 48 \cdot 47 \cdot 46}{52!} = \binom{49}{4} \Big/ \binom{52}{4}.$$

5. (De Méré) 掷 4 次骰子至少得到一个 6 的概率与掷两个骰子 24 次至少得一次双 6 的概率哪个大？

   ▶(电脑计算)
   $$1 - (5/6)^4 > 1 - (35/36)^{24}.$$

6. 甲掷 $6n$ 个骰子要得到至少 n 个 6, 乙掷 $6(n+1)$ 个骰子要得到至少 $n+1$ 个 6, 问谁的概率大？

7. 在两付牌中取 26 张至少有 2 张牌完全一样的概率是多少？

   ▶
   $$2^{26} \cdot \binom{52}{26} \Big/ \binom{104}{26}.$$

8. 一个盒子里有红球与黑球，现在任取两个球都是红球的概率是 $\frac{1}{2}$. 问 (1) 盒子中至少有几个球？ (2) 已知有偶数个黑球，盒子中至少该有几个球？

   ▶设其中 a 个红球， b 个黑球，则
   $$\frac{1}{2} = \frac{a(a-1)}{(a+b)(a+b-1)}.$$
   (1) $a = 3$, $b = 1$; (2) $a = 15$, $b = 6$.

9. 父亲为了鼓励儿子打网球宣称如果他能赢得三场与父亲和教练的比赛中连续的两场，他将获得一笔奖金．他可以选择比赛的顺序为 (1) 父亲 – 教练 – 父亲； (2) 教练 – 父亲 – 教练．教练比父亲打得好．问为了增加获得奖金的机会，他应该选择哪个顺序？

   ▶(1) 的概率小于 (2) 的概率.

10. 假设某人在赌场使用下面的策略：先押 1 元钱，赢了就离开，输了就再赌一次，押 2 元钱，然后不论输赢都离开．设每次输赢都是等可能的．求他最后赢钱的概率．问这个策略好不好？

11. 从数 $1, 2, \cdots, n$ 中任取 k 个，求其中没有两个数相邻的概率.
▶$\binom{n-k+1}{k}/\binom{n}{k}$. $1 \le j_1 < j_2 < \cdots < j_k \le n$ 且 $j_i - j_{i-1} > 1$, 做可逆的变换 $l_u = j_u - (u-1)$, 那么 $1 \le l_1 < l_2 < \cdots < l_k \le n-(k-1)$ 且 $l_i - l_{i-1} > 0$.

12. 另一个配对问题，考虑 n 对夫妇围坐于圆桌，且任何人的两边都是异性，求没有一对夫妇坐在一起的概率.
▶不妨给坐位依次编号 $1, 2, \cdots, 2n$. 坐位 1 坐一个女人. 事件 A_i: i 与 $i+1$ 是夫妻 ($2n+1$ 看成 1).

$$\begin{aligned}\mathsf{P}(\bigcap_{i=1}^{2n} A_i^c) &= 1 - \mathsf{P}(\bigcup_{i=1}^{2n} A_i) \\ &= 1 - \sum_{k=1}^{2n} (-1)^{k-1} \sum_{i_1 < \cdots < i_k} \mathsf{P}(A_{i_1} \cap \cdots \cap A_{i_k}).\end{aligned}$$

计算

$$\mathsf{P}(A_i) = \frac{n \cdot (n-1)!^2}{n!^2} = \frac{1}{n};$$
$$\mathsf{P}(A_i \cap A_j) = \begin{cases} n(n-1) \cdot (n-2!^2/n!^2 = \frac{1}{n(n-1)}, & j-i > 1, \\ 0, & j-i = 1. \end{cases}$$
$$\cdots\cdots$$
$$\mathsf{P}(A_{i_1} \cap \cdots A_{i_k}) = \begin{cases} \frac{1}{(n)_k}, & i_u - i_{u-1} > 1,\ 1 \le u \le k-1, \\ 0, & \text{otherwise}. \end{cases}$$

现在要算从 $2n$ 个圆桌坐位取 k 对不相交相邻坐位的方法数 $S_{k,n}$. 先算从 m 个排坐位中取 k 对不相交相邻坐位的方法数 $N_{k,m}$, 参考 11 题， $N_{k,m} = \binom{m-k}{k}$. 算 $S_{k,n}$ 时分两种情况：取 $\{1, 2n\}$ 或不取. 因此

$$S_{k,n} = N_{k-1,2n-2} + N_{k,2n} = \binom{2n-k}{k} \cdot \frac{2n}{2n-k}.$$

带入上面的和式，所求概率为

$$\sum_{k=0}^{n} (-1)^k \binom{2n-k}{k} \cdot \frac{2n}{2n-k} \cdot \frac{1}{(n)_k}.$$

13. 一个盒子中有 m 个白球和 n 个黑球，不放回地每次取一个一直到第 r 个黑球取出. 求这时取出的总的球个数是 k 的概率.

▶
$$\mathsf{P}(A_k)=\frac{\binom{k-1}{r-1}(m)_{k-r}(n)_r}{(m+n)_k}$$

14. (1) 从 n 双不同的鞋子中任取 $2k$ 只，求其中恰有 s 双的概率. (2) 从 n 双相同的鞋子中任取 $2k$ 只，求其中恰有 s 双的概率.

▶(1)
$$\frac{\binom{n}{s}\binom{n-s}{2k-2s}\cdot 2^{2k-s}}{\binom{2n}{2k}}$$

(2)
$$\frac{2\binom{n}{s}\binom{n-s}{2k-2s}}{\binom{2n}{2k}}$$

15. 掷三次骰子，(a) 求所得的点数一次比一次大的概率. (b) 求点数依次不减小的概率.

▶(1) $\binom{6}{3}/6^3$; (2) $\binom{8}{3}/6^3$, n 个数中可重复取 k 个, 依次不减. $1\le j_1\le j_2\cdots\le j_k\le n$, 做可逆变换 $l_u=j_u+(u-1)$, 这相当于从 $n+k-1$ 个数中可重复地取 k 个，依次严格递增.

16. (a) 两人各掷两枚硬币，求他们所得正面数一样的概率； (b) 两人各掷三个骰子，求他们所得的点数一样的概率.

17. 将 n 条同样的绳子的 $2n$ 个头任意两两连接，求恰好连成一个环的概率.

▶共 $2n$ 个线头，编号排好，有 $(2n)!$ 种排法. 要结成一个圈，线头不能和另一头相连，先取一个线头， $2n$ 种选法，然后它只能与其他 $n-1$ 根线的任一头连接，有 $2n-2$ 种选法，接着把两根连在一起的线看成一根继续，因此共有

$$2n(2n-2)(2n-2)(2n-4)\cdots 6\cdot 4\cdot 4\cdot 2\cdot 2=(2n)!!(2n-2)!!$$

种选法，概率为
$$\frac{(2n)!!(2n-2)!!}{(2n)!}=\frac{(2n-2)!!}{(2n-1)!!}.$$

18. 一个袋中有 A 个球，其中 a 个白球，其他是黑球，依次一个一个地不放回取球，问第 k 次取球时是首次取得白球的概率是多少？

19. 三人 A, B, C 以此顺序连续地掷骰子. 求： (1) A 第一个掷出 6 点， B 第二个， C 第三个的概率. (2) 求第一个 6 点是 A 掷出的，第二个 6 点是 B 掷出的，第三个 6 点是 C 掷出的概率.

►(应该放在 §1.4) 分别用 A_1, A_2 表示 (1), (2) 中叙述的两个事件. (1) 不管 A, 那么在这个比赛中 B 比 C 先得 6 点的概率是 $p = 1/6 + (5/6)^2 \cdot 1/6 + (5/6)^4 \cdot 1/6 + \cdots = 6/11$. 因此

$$\mathsf{P}(A_1) = 1/6p + (5/6)^3 \cdot 1/6 \cdot p + (5/6)^6 \cdot 1/6 \cdot p + \cdots = 6/11 \cdot 36/91.$$

(2) 由乘法公式

$$\begin{aligned}\mathsf{P}(A_2) &= \mathsf{P}(\text{ABC..., A 先得 6})\mathsf{P}(\text{BCA..., B 先得 6})\mathsf{P}(\text{CAB..., C 先得 6})\\ &= (\mathsf{P}(\text{ABC...., A 先掷得 6}))^3\\ &= \left(\sum_{n\geq 0}\left(\frac{5}{6}\right)^{3n}\cdot\frac{1}{6}\right)^3 = \left(\frac{36}{91}\right)^3\end{aligned}$$

20. 任取一个自然数，在附录 ?? 的意义下，求下列事件的概率. (1) 此数平方的个位数等于 1. (2) 此数四次方后的个位数等于 1.

►(1) 2/10; (2) 4/10

21. 设 A, B, C 是事件，如果三个事件不可能同时发生，也不可能其中一个事件单独发生，证明：

$$\mathsf{P}(A \cup B \cup C) = \frac{1}{2}(\mathsf{P}(A) + \mathsf{P}(B) + \mathsf{P}(C)).$$

►不能同时发生： $A \cap B \cap C = \emptyset$, 不能单独发生：

$$A \cap B^c \cap C^c = A^c \cap B \cap C^c = A^c \cap B^c \cap C = \emptyset.$$

因此 $A \subset B \cup C$, $B \subset C \cup A$, $C \subset A \cup B$, 那么

$$\mathsf{P}(A) = \mathsf{P}(AB \cup AC) = \mathsf{P}(AB) + \mathsf{P}(AC) - \mathsf{P}(ABAC) = \mathsf{P}(AB) + \mathsf{P}(AC).$$

类似地， $\mathsf{P}(B) = \mathsf{P}(BA) + \mathsf{P}(BC)$, $\mathsf{P}(C) = \mathsf{P}(CA) + \mathsf{P}(CB)$, 最后

$$\begin{aligned}\mathsf{P}(A \cup B \cup C) &= \mathsf{P}(A) + \mathsf{P}(B) + \mathsf{P}(C) - \mathsf{P}(A \cap B) - \mathsf{P}(B \cap C) - \mathsf{P}(C \cap A)\\ &= \frac{1}{2}(\mathsf{P}(A) + \mathsf{P}(B) + \mathsf{P}(C)).\end{aligned}$$

22. 从 n 阶行列式的一般展开式中任取一项，问这项中包含主对角线元素的概率是多少？

23. 一个人试图扔一个一元硬币 (直径 3cm) 到 2 米外的一个边长为 4cm 的正方形台面上，现在他的硬币已经落在台面上，问硬币完全在台面里的概率是多少？

24. 向一个画着等距离平行线的平面上任意地投一个边长为 a,b,c 的三角形，平行线间的距离为 l, 其中 $a,b,c<l$, 求三角形与平行线相交的概率.
   ►利用 21 题和 Buffon 投针问题.
25. 回到配对问题 (例 ??), 用 B_k 表示 k 个人取自己的帽子而没有人选中自己帽子 这个事件，利用组合计数直接证明：当 $k\geq 3$ 时,

$$|B_k|=(k-1)(|B_{k-1}|+|B_{k-2}|),$$

   且 $|B_1|=0, |B_2|=1$. 由此求出 $|B_k|$.

§1.3 概率空间与随机变量

习 题

1. 证明： $1_{A\cup B}=1_A\vee 1_B$, $1_{A\cap B}=1_A\wedge 1_B$, $1_{A^c}=1-1_A$, $1_{A\Delta B}=|1_A-1_B|$.
2. 证明： $1_{A\cup B}=1_A+1_B$ 当且仅当 A,B 不相交.
3. 证明：任意多个 σ- 域的交仍然是 σ- 域.
4. 如果一个 σ- 域的元素个数有限，证明：其元素个数一定是 2 的整数幂.
5. 设 $(\Omega,\mathscr{F})$ 是可测空间， P 是 $\mathscr{F}$ 上非负函数且 $\mathsf{P}(\Omega)=1$, 证明下列命题等价.
   (a) P 是可列可加的;
   (b) P 是有限可加的且是下连续的;
   (c) P 是有限可加的且次可列可加的.
6. 设 A,B 是概率空间中的两个事件且 $\mathsf{P}(A)=1$, 证明： $\mathsf{P}(A\cap B)=\mathsf{P}(B)$.
7. 设 P 是概率， $\mathsf{P}(A)=\frac{3}{4}$, $\mathsf{P}(B)=\frac{1}{3}$, 证明： $\frac{1}{12}\leq \mathsf{P}(A\cap B)\leq \frac{1}{3}$.
8. 证明 (1) 如果 Ω 是可列无限的，那么在 Ω 上不存在使所有基本事件等可能出现的概率； (2) Ω 上的事件域 $\mathscr{F}$ 离散当且仅当 $\mathscr{F}$ 上的任何随机变量是离散的.
9. 设有 m 持 50 元钱的人和 n 个持 100 元钱的人在一个窗口排队买票， $m\geq n$, 票价是 50 元且窗口开始没有零钱. 求所有买票的人都不需要等待找钱的概率.
   ►利用反射原理，用 S_i 表示前 i 个人中持 50 元钞票的人数减去持 100 元钞票的人数. 则买票期间不需要等待等价于一条从 $(0,0)$ 最终到 $(m+n,m-n)$ 且在 x 轴上方 (可接触此轴) 的格点轨道 $\{(i,S_i)\}$.

10. 证明：

$$\mathsf{P}(\bigcap_{i=1}^n A_i) = \sum_i \mathsf{P}(A_i) + (-1)\sum_{i<j}\mathsf{P}(A_i \cup A_j) + \sum_{i<j<k}\mathsf{P}(A_i \cup A_j \cup A_k)$$
$$+ \cdots + (-1)^{n-1}\mathsf{P}(A_1 \cup \cdots \cup A_n).$$

►De Morgan 对偶，定理 1.2.1

11. 证明：

$$\mathsf{P}(\bigcup_{i=1}^n A_i) \geq \sum_{i=1}^n \mathsf{P}(A_i) - \sum_{i<j}\mathsf{P}(A_i \cap A_j).$$

►数学归纳法，定理 1.2.1

12. 证明：

$$\mathsf{P}(\bigcup_{i=1}^n A_i) \leq \min_k \left\{ \sum_{i=1}^n \mathsf{P}(A_i) - \sum_{i:i\neq k}\mathsf{P}(A_i \cap A_k) \right\}.$$

►数学归纳法

13. 设有事件 $A_1, \cdots, A_n$, N_k 表示恰有 k 个 A_i 发生的事件．证明：

$$\mathsf{P}(N_k) = \sum_{i=0}^{n-k}(-1)^i \binom{k+i}{k} S_{k+i},$$

其中 $S_j = \sum_{i_1<i_2<\cdots<i_j}\mathsf{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_j})$.
►$[n] := \{1, 2, \cdots, n\}$. N_k 表示 $A_1, \cdots, A_n$ 中恰有 k 个发生．则

$$N_k = \bigcup_{I\subset[n],|I|=k} \left((\bigcap_{i\in I} A_i) \cap (\bigcup_{i\notin I} A_i)^c \right).$$

令 $A_I := \bigcap_{i\in I} A_i$, 那么 $N_k = \bigcup_I (A_I \cap (\cup_{i\notin I} A_i)^c)$, 然后应用定理 1.2.1.

14. 设 ξ 是随机变量，分布函数是 F, 证明：对任何 $x \in \mathbf{R}$, $\mathsf{P}(\xi < x) = F(x-)$, x 处的左极限．从而证明 F 连续当且仅当对任何 $x \in \mathbf{R}$, $\mathsf{P}(\xi = x) = 0$, 这时也说 ξ 是散布的 (diffused).

15. 设有事件 $A_1, \cdots, A_n$, 其中至少有一个一定发生，但不可能有多于两个同时发生， $\mathsf{P}(A_i) = p$, $\mathsf{P}(A_i \cap A_j) = q$, $i \neq j$, 证明： $p \geq \frac{1}{n}$, $q \leq \frac{2}{n}$.
►依条件 $\Omega = \bigcup_i A_i$, 故 $1 = \mathsf{P}(\bigcup_i A_i) \leq np$, 即 $p \geq 1/n$. 再由于不可能有多于两个事件同时发生，故 $1 = np - n(n-1)/2 \cdot q$, 因此

$$q = \frac{2(np-1)}{n(n-1)} \leq 2/n.$$

16. 实数 m 称为随机变量 ξ 的中点，如果 $\mathsf{P}(\xi < m) \le \frac{1}{2} \le \mathsf{P}(\xi \le m)$. 证明：任何随机变量至少有一个中点，且中点的集合是一个闭区间.
►存在 x 使得 $\mathsf{P}(\xi \le x) \ge 1/2$. 令 $m := \inf\{x : \mathsf{P}(\xi \le x) \ge 1/2$, 那么可以证明 $F(m) \ge 1/2 \ge F(m-)$.
17. 设 ξ 是随机变量, 问是否存在正实数 a, 使得 ξ 与 $\xi + a$ 有相同的分布函数？
►不可能. 如果有的话，那么 ξ 的分布函数满足 $F(x) = F(x-a)$ 对所有 x 成立，也就是说 F 是周期函数，这是不可能的.
18. 设 ξ, η 是两个随机变量， I 是个区间，证明：
$$|\mathsf{P}(\xi \in I) - \mathsf{P}(\eta \in I)| \le \mathsf{P}(\xi \ne \eta).$$
►
$$\begin{aligned}\mathsf{P}(\xi \in I) &= \mathsf{P}(\xi \in I, \xi = \eta) + \mathsf{P}(\xi \in I, \xi \ne \eta) \\ &\le \mathsf{P}(\eta \in I, \xi = \eta) + \mathsf{P}(\xi \ne \eta) \\ &\le \mathsf{P}(\eta \in I) + \mathsf{P}(\xi \ne \eta).\end{aligned}$$
19. 证明：一个连续的分布函数是一致连续的.

§1.4 条件概率与全概率公式

习 题

1. 掷一个骰子 n 次， A_{ij} 表示第 i 次与 j 次得到同样点数的事件，证明： $\{A_{ij} : 1 \le i, j \le n\}$ 两两独立，但不是相互独立的.
2. 设 p 是素数，从 $\Omega = \{1, 2, \cdots, p\}$ 中随机取个数， $A, B \subset \Omega$, 事件 A, B 表示所取数分别落在 A, B 集内. 若 A, B 独立，那么 A, B 中至少有一个是 $\varnothing$ 或者 Ω.
3. 掷一枚均匀硬币 100 次， ξ 表示正面次数，问 $\mathsf{P}(48 \le \xi \le 52)$ 与 $\mathsf{P}(\xi \ge 60)$ 哪个大？给出确切理由.
4. 设 $A_1, A_2, \cdots, A_n, \cdots$ 是独立事件序列， $A := \limsup A_n$, 证明： $\mathsf{P}(A)$ 是 0 或者 1.
►要证明 A 与自身独立. 分两步 (1) 证明对任何 n, $A_1, A_2, \cdots, A_n, \bigcup_{k>n} A_k$ 独立. 因为对任何 $m > 0, A_1, \cdots, A_n, \cdots, A_{n+m}$ 独立,故 $A_1, \cdots, A_n, \bigcup_{k=n+1}^{n+m} A_k$ 独立. 由概率的连续性推出结论. (2) 对任何 n,
$$A = \bigcap_{m\ge 1} \bigcup_{k \ge m} A_k = \bigcap_{m > n} \bigcup_{k \ge m} A_k.$$

因此 $A, A_1, \cdots, A_n$ 独立．而 n 是任意的，故 $A, A_1, \cdots, A_n, \cdots$ 独立．推出 A 与自身独立．

上面的方法，本质上要用到 Dynkin 引理，不是一个好方法．另外一个简单方法，由 07 级陈云霄提供：令 $B_n = \bigcap_{k\geq n} A_k^c$, 那么

$$\mathsf{P}(A) = \lim_n \mathsf{P}(\bigcup_{k\geq n} A_k) = 1 - \lim_n \mathsf{P}(B_n),$$

B_n 递增且由独立性得 $\mathsf{P}(B_n) = \prod_{k\geq n} \mathsf{P}(A_k^c)$. 设 $a = \lim_n \mathsf{P}(B_n)$, 反设 $0 < a < 1$, 那么存在 m, 当 $n \geq m$ 时 有 $a(1+a)/2 \leq \mathsf{P}(B_n) \leq a$,

$$\prod_{m\leq k<n} \mathsf{P}(A_k^c) = \frac{\mathsf{P}(B_m)}{\mathsf{P}(B_n)} \geq \frac{1+a}{2}, \tag{1.4.1}$$

让 n 趋于无穷得 $\mathsf{P}(B_m) \geq \frac{1+a}{2} > a$, 矛盾.

受上述式子的启发，我们只需要在 (1.4.1) 的等号 (对任何 m 成立) 两端让 $n \longrightarrow \infty$, 得 $\mathsf{P}(B_m) = a\mathsf{P}(B_m)$, 因此 $a = 1$ 或者对任何 m, $\mathsf{P}(B_m) = 0$, 即 $a = 0$.

5. 掷两个骰子，证明：和为 7 的事件与第一个骰子的点数是独立的.
6. 事件 A 与所有事件独立当且仅当 $\mathsf{P}(A)$ 是 0 或者 1.
7. 设 ξ, η 是独立同分布随机变量，取值 x 的概率是 2^{-x}, $x = 1, 2, \cdots$. 求：$\mathsf{P}(\min(\xi,\eta) \leq x)$, $\mathsf{P}(\eta > \xi)$, $\mathsf{P}(\xi = \eta)$, $\mathsf{P}(\xi \geq k\eta)$, (其中 k 是正整数,) $\mathsf{P}(\xi$ 整除 $\eta)$, $\mathsf{P}(\xi = r\eta)$, 其中 r 是实数.
8. 设 X_1, X_2, X_3 是独立随机变量， $\mathsf{P}(X_i = x) = (1-p_i)p_i^{x-1}$, $x = 1, 2, \cdots$, $i = 1, 2, 3$.
   (a) 求 $\mathsf{P}(X_1 < X_2 < X_3)$.
   (b) 求 $\mathsf{P}(X_1 \leq X_2 \leq X_3)$.
9. 如果 X 是参数为 n, p 的二项分布， Y 是 m, p 的二项分布，且 X, Y 独立. 算 $X + Y$ 的分布.
10. 设 X 是非负整数值随机变量，对任何整数 $k \geq 0$ 定义 $h(k) := \mathsf{P}(X = k | X \geq k)$. 如果 $\{U_i : i \geq 0\}$ 是独立且都服从 $[0, 1]$ 上均匀分布，证明：$Z := \min\{n : U_n \leq h(n)\}$ 与 X 同分布.
    ►对任何 $k \geq 0$, $Z = k$ 等价于当 $i < k$ 时 $U_i > h(i)$ 而 $U_k \geq h(k)$. 然后利用独立性和均匀分布的性质.
11. 证明：事件 $A_1, \cdots, A_n$ 独立当且仅当指标 $1_{A_1}, \cdots, 1_{A_n}$ 独立.
12. 重复 Bernoulli 试验，求第 n 次成功发生在第 m 次失败前的概率.

►用 ξ 表示第 n 成功发生时的总试验次数，它是 Pascal 分布的．所求概率是 $\mathsf{P}(\xi < n+m)$.

13. 证明：(1) ξ 是几何分布的当且仅当存在 $q \in (0,1)$ 使得对任何 $n \geq 0$, $\mathsf{P}(\xi > n) = q^n$; (2) 几何分布有遗忘性：对任何 $m, n \in \mathbf{Z}_+$, 有

$$\mathsf{P}(\xi > m+n | \xi > m) = \mathsf{P}(\xi > n).$$

(3) 有遗忘性的 $\mathbf{Z}_+$ 值随机变量的分布一定是几何分布.

14. 如果随机变量 ξ, η 是独立的且都服从几何分布, 那么 $\xi \wedge \eta$ 也服从几何分布.

15. 设 u_n 是掷 n 次硬币其中没有连续的 4 次正面的概率, 那么对 $n \geq 4$, 有公式

$$u_n = \frac{1}{2}u_{n-1} + \frac{1}{4}u_{n-2} + \frac{1}{8}u_{n-3} + \frac{1}{16}u_{n-4},$$

以此计算 u_8.

►全概率公式，并递推得 $u_8 = 26/32$.

16. 有标号为 $1, 2, \cdots, m$ 的 m 张卡片，随机地一张一张抽取，已知第 k 张是前 k 张中最大的，问它是 m 的概率.

17. 设有 A, B 两个盒子，分别有三个白球和三个黑球，每次从两盒子中任取一个球交换，求 n 次后盒子中的球的颜色仍然是相同的概率.

►全概率公式得 $p_n = \frac{1}{9}(1 - p_{n-1})$, $p_0 = 1$, $p_1 = 0$.

18. 设 A 有三个白球，B 盒有三个白球，一个黑球，每次从两盒子中任取一个球交换，求 n 次后黑球仍然在 B 盒的概率.

►全概率公式得

$$p_n = p_{n-1} \cdot \frac{3}{4} + (1 - p_{n-1}) \cdot \frac{1}{3},$$

$p_0 = 1$.

19. 设 A, B 是一随机试验的两个互斥事件，重复这个试验，问 B 事件比 A 事件先出现的概率是多少？

►用 ξ, η 分别表示 B, A 首次出现时的试验次数. 那么所求概率是 $\mathsf{P}(\xi < \eta) = \sum_{k \geq 1} \mathsf{P}(\xi = k, \eta > k)$. 而 $\xi = k, \eta > k$ 等价于前 $k-1$ 次试验两者都没有出现，而第 k 次试验 B 出现．因此 $\mathsf{P}(\xi = k, \eta > k) = (1 - \mathsf{P}(A \cup B))^{k-1}\mathsf{P}(B)$, 故而 $\mathsf{P}(\xi > \eta) = \mathsf{P}(B)/(\mathsf{P}(A) + \mathsf{P}(B))$.

20. 美洲有一种骰子游戏，赌徒掷两个骰子．如果掷出 7 或 11, 他就赢了．如果掷出 2, 3 或 12, 他就输了．如果掷出其他，记下这个点，再掷骰子一直到掷出这个点数，那么他赢，或者掷出 7 点，那么他输，问赌徒赢的概率是多少？

▶首先写出骰子点数和 ξ 的分布律. 那么由全概率公式

$$\mathsf{P}(A)=\sum_k \mathsf{P}(A|\xi=k)\mathsf{P}(\xi=k)=6/36+2/36+\sum_{k\in\{4,5,6,8,9,10\}}\mathsf{P}(A|\xi=k)\mathsf{P}(\xi=k),$$

而 $\mathsf{P}(A|\xi=k)=\mathsf{P}($ 在掷出 7 前先掷出 k 点 $)$. 由 18 题,

$$\mathsf{P}(A|\xi=4)=\mathsf{P}(A|\xi=10)=3/9,$$
$$\mathsf{P}(A|\xi=5)=\mathsf{P}(A|\xi=9)=4/10,$$
$$\mathsf{P}(A|\xi=6)=\mathsf{P}(A|\xi=8)=5/11.$$

最后,

$$\mathsf{P}(A)=244/495=\frac{1}{2}-\frac{7}{990}.$$

几乎 50%, 设计得非常精确.

21. 设随机变量 X,Y,ξ 独立，且 X,Y 的分布函数分别是 F,G,ξ 服从参数 p 的 Bernoulli 分布，写出 $\xi X+(1-\xi)Y$ 的分布函数.
▶$pF+(1-p)G$
22. 从 $\{1,\cdots,n\}$ 中独立重复地取一个数， ξ_k 是第 k 次取得的数. 令 τ 是首次拿到偶数的时间，即
$$\tau=\inf\{k:\xi_k \text{ 是个偶数}\}.$$
求复合随机变量 ξ_τ 的分布. 注意： ξ_τ 的严格定义 $\xi_\tau:=\sum_{k\geq 1}\xi_k\cdot 1_{\{\tau=k\}}$.
23. 如果 $\{A_n:n\geq 1\}$ 独立，证明：
$$\mathsf{P}\left(\bigcap_{n\geq 1}A_n\right)=\prod_{n\geq 1}\mathsf{P}(A_n).$$
24. 一个盒子中 5 个球， 3 黑 2 白，两人依次取摸球不返回，记 A 为第一个人取得黑球的事件， B 为第二个人取得黑球的事件，求 $\mathsf{P}(A|B)$ 与 $\mathsf{P}(B|A)$.
25. 掷一个骰子点数为 X, 然后掷 X 枚硬币，记硬币正面数为 Y. 求 Y 的分布.
26. 设 ξ_1,ξ_2 独立且分别服从参数为 λ_1,λ_2 的 Poisson 分布，对整数 $n\geq k\geq 0$, 求 $\mathsf{P}(\xi_1=k|\xi_1+\xi_2=n)$.
27. 设 $\xi_1,\cdots,\xi_r$ 是独立同分布随机变量, 都服从几何分布, 证明在 $\xi_1+\cdots+\xi_r=n$ 的条件下， $(\xi_1,\cdots,\xi_r)$ 的分布是均匀分布，即
$$\mathsf{P}(\xi_1=n_1,\cdots,\xi_r=n_r|\xi_1+\cdots+\xi_r=n)=1/\binom{n-1}{n-r}.$$

28. 在 n 次重复 Bernoulli 试验中，ξ_i 是第 i 次试验中成功的指标 $(0 \le i \le n)$. 求在 $\xi_1 + \cdots + \xi_n = r, r \le n$ 的条件下，ξ_i 的分布律.
29. 设事件 A, B 独立且 $A \subset B$, 证明：$\mathsf{P}(A) = 0$ 或者 $\mathsf{P}(B) = 1$.
30. 设 ξ, η 是独立 Bernoulli 分布的，成功概率为 p, 定义 ζ 为 $\xi + \eta$ 为偶数这个事件的指标，问何时 ξ 与 ζ 独立？
    ►$p = 0, 1, 1/2$.
31. 设 ξ, η 独立且是 $1, -1$ 上等可能分布的，$\zeta := \xi\eta$. 证明：ξ, η, ζ 两两独立但不独立.

§1.5 随机变量的期望

习 题

1. 设 ξ 是离散随机变量，A 是事件，那么 $\mathsf{E}(\xi; A) = \sum_{x \in R(\xi)} x\mathsf{P}(\xi = x, A)$.
2. 平均地讲，掷一个硬币多少次就可以得到连续的两次正面或连续的两次反面？
   ►ξ: 首次出现两次连续正面或反面的试验数. 那么 $\xi \ge 2$. 显然 $\mathsf{P}(\xi = 2) = 1/2$, $\mathsf{P}(\xi = 3) = 1/4$, $\mathsf{P}(\xi = k) = 2/2^k$. 因此 $\mathsf{E}\xi = 3$.
3. 设 ξ 为非负整值随机变量，其数学期望存在，证明：$\mathsf{E}\xi = \sum_{n=1}^{\infty} \mathsf{P}(\xi \ge n)$.
   ►
$$\begin{aligned} \mathsf{E}\xi &= \sum_{n \ge 1} n\mathsf{P}(\xi = n) \\ &= \sum_{n \ge 1} \sum_{k=1}^{n} \mathsf{P}(\xi = n) \\ &= \sum_{k \ge 1} \sum_{n \ge k} \mathsf{P}(\xi = n) = \sum_{k \ge 1} \mathsf{P}(\xi \ge k). \end{aligned}$$
4. 设 ξ 服从参数为 n, p 的二项分布，证明：
$$\mathsf{E}\frac{1}{1+\xi} = \frac{1-(1-p)^{n+1}}{(n+1)p}.$$
   ►$\mathsf{E}(1/(1+\xi)) = (1 - q^{n+1})/p^{n+1}$.
5. 设 $\{\xi_n : n \ge 0\}$ 成功概率为 p 的独立 Bernoulli 随机序列，定义
$$X = \inf\{n \ge 1 : \xi_n > \xi_{n-1},$$

也就是说 X 是首次从失败到成功的转折，求 $\mathsf{E}X$.

►相当于 01 出现时刻. 先等 0 出现，期望为 $1/q-1$ (注意序列 ξ_n 从零开始)，再等 1 出现，期望为 $1/p$, 因此答案为 $1/p+1/q-1$.

6. (Coupon 问题) 推销装有卡通人物的卡片的某种方便面，一套卡片共 N 张. 在每包方便面中随机地放置一张卡通卡片. 求买了 n 包方便面仍然没有收集到一整套卡通卡片的概率，与收集到一套卡片时所买方便面的包数的期望.

   ►等同于放球问题：往 N 个盒子里放球，用 ξ 表示首次没有空盒时所放的球数. $\xi>n$ 表示投放 n 个球后仍然有空盒. 用 B_i 表示第 i 个盒子空. 那么由定理 1.2.1,

$$\mathsf{P}(\xi>n)=\mathsf{P}\left(\bigcup_{i=1}^N B_i\right)=\sum_{k=1}^N(-1)^{k-1}\binom{N}{k}\mathsf{P}(\bigcap_{i=1}^k B_i)=\sum_{k=1}^N(-1)^{k-1}\binom{N}{k}\cdot\frac{(N-k)^n}{N^n}.$$

   那么

$$\mathsf{E}\xi=\sum_{n\geq 0}\mathsf{P}(\xi>n)=\sum_{k=1}^N(-1)^{k-1}\binom{N}{k}\cdot\frac{N}{k}.$$

   ►另外有个直接的方法可以算期望. 令 $\eta_1=1$, η_i 是从投放 η_{i-1} 个球之后算起，等待第 i 个非空盒子出现时投放的球数. 那么 η_i 实际上是服从概率为 $(N-(i-1))/N$ 的几何分布. 而 $\xi=\sum_{i=1}^N\eta_i$, 因此 $\mathsf{E}\xi=\sum_{i=1}^N N/(N-i+1)=\sum_{k=1}^N N/k$.

   ►这样，我们证明了一个恒等式

$$\sum_{k=1}^N(-1)^{k-1}\binom{N}{k}\cdot\frac{N}{k}=\sum_{k=1}^N\frac{N}{k}.$$

7. 某人用 n 把外形相似的钥匙去开门，只有一把能打开. 今逐个任取一把试开，直至打开门为止. 分别考虑每次试毕 (1) 不放回； (2) 放回两情形. 求试开次数 ξ 的数学期望.

   ►在放回情况下， ξ 服从参数为 $1/n$ 的几何分布. 在不放回的情况下，服从 $\{1,\cdots,n\}$ 上的均匀分布. 因此期望分布是 n 和 $(n+1)/2$.

8. 如果 X 服从参数为 λ 的 Poisson 分布，证明：

$$\mathsf{E}(X(X-1)\cdots(X-k))=\lambda^{k+1},\ k=0,1,2\cdots.$$

   求 X 的方差.

9. 从数字 $1,\cdots,n$ 中任取两个不同的数，求两数之差绝对值的期望当 n 趋于无穷时的极限.

10. 8 个男单身与 7 女单身随机坐在一排 15 个座位上，ξ 表示其中形成的相邻且有可能成婚的对数，(如用 M,F 分别表示男与女，例如 MMMFFFMFM-MMMFFF 的坐次，$\xi = 5$.) 求 $\mathsf{E}\xi$. (提示：用 ξ_1 表示头两个座位的两个人是异性这个事件的指标. 求 $\mathsf{P}(\xi_1 = 1)$.)

   ►用 ξ_i 表示 i 与 $i+1$ 座两人是异性这个事件的指标，$1 \leq 14$. 那么 $\mathsf{P}(\xi_i = 1) = 2 \cdot 7 \cdot 8 \cdot 13!/15! = 8/15$. 而 $\mathsf{E}\xi = \mathsf{E}\xi_1 + \cdot + \mathsf{E}\xi_{14} = 14 \cdot 8/15$.

11. 设 ξ, η 是离散随机变量，证明：(1) ξ, η 独立当且仅当对任何有界函数 (或非负函数) f, g, 有 $\mathsf{E}f(\xi)g(\eta) = \mathsf{E}f(\xi)\mathsf{E}g(\eta)$. (2) 如果 ξ, η 独立，则对任何函数 f, g, 有 $f(\xi), g(\eta)$ 也独立.

12. 设 ξ, η 是二值 (值域仅有两个元素) 随机变量，证明：ξ, η 独立当且仅当它们不相关.

   ►不妨设 ξ, η 分别是事件 A, B 的指标. 那么 $\mathsf{E}\xi\eta = \mathsf{P}(A \cap B)$, $\mathsf{E}\xi = \mathsf{P}(A)$, $\mathsf{E}\eta = \mathsf{P}(B)$, 因此不相干等价于 $\mathsf{P}(A \cap B) = \mathsf{P}(A)\mathsf{P}(B)$, 即 A, B 独立，这等价于 ξ, η 独立.

13. 在一个球面上 10% 涂上兰色，其他红色，证明：不管色彩怎么涂，都可以内接一个正方体，其所有顶点都是红色的.

   ►用 ξ_i 表示第 i 顶点是红色这个事件的指标. 那么 $\mathsf{E}\sum_{i=1}^{8} \xi_i = 8 \cdot 9/10 > 7$. 因此 $\mathsf{P}(\sum_{i=1}^{8} \xi_i \geq 8) > 0$.

14. 掷一个骰子到所有点数都至少出现一次即停止，ξ 表示掷的次数. 求 $\mathsf{E}\xi$.

15. 一个盒子里有标号为 $1, 2, \cdots, n$ 的 n 个球，不放回地随机取 k 个，求它们的和的期望与方差.

   ►ξ_i 表示第 i 次取出的数. 那么 $\mathsf{E}(\xi_1 + \cdots + \xi_k) = k \cdot (n+1)/2$.

16. 袋子 A 里有 n 个红球，袋子 B 里有 n 个兰球，每次都从两个袋子里各取一球交换，证明：在 k 次交换后 A 袋中红球数的期望为 $\frac{1}{2}n(1 + (1 - \frac{2}{n})^k)$.

   ►用 ξ_k 表示 k 次交换后 A 袋中的红球数. 那么在 $\xi_{k-1} = m$ 的条件下，ξ_k 以概率 $(1 - m/n)^2$, $(m/n)^2$, $1 - (1 - m/n)^2 - (m/n)^2$ 分布在 $m+1$, $m-1$, m 上. 因此

$$
\begin{aligned}
&\mathsf{E}(\xi_k | \xi_{k-1} = m) \\
&\quad = (m+1) \cdot (1 - m/n)^2 + (m-1) \cdot (m/n)^2 + m \cdot (1 - (1 - m/n)^2 - (m/n)^2) \\
&\quad = (1 - 2/n)m + 1.
\end{aligned}
$$

   因此 $\mathsf{E}\xi_k = \sum_m \mathsf{E}(\xi_k | \xi_{k-1} = m) \cdot \mathsf{P}(\xi_{k-1} = m) = 1 + (1 - 2/n)\mathsf{E}\xi_{k-1}$. 然后递推且利用初值 $\xi_0 = n$.

17. 连续地掷一个正面概率为 p 的硬币，X_n 是得到 n 次连续正面所掷的次数，

证明： $\mathsf{E}X_n = \sum_{k=1}^{n} p^{-k}$.

►η 为反面首次出现的时间． $\mathsf{P}(\eta > n) = p^n$.

$$\begin{aligned}\mathsf{E}X_n &= \sum_{k=1}^{n} \mathsf{E}(X_n|\eta = k)\mathsf{P}(\eta = k) + \mathsf{E}(X_n|\eta > n)\mathsf{P}(\eta > n) \\ &= \sum_{k=1}^{n} qp^{k-1} \cdot (k + \mathsf{E}X_n) + n \cdot p^n \\ &= \sum_{k=1}^{n} p^{k-1} + (1 - p^n) \cdot \mathsf{E}X_n.\end{aligned}$$

18. 设有涂有红与两种颜色的圆周，如果其任意内接正三角形总有一个顶点是红色，证明：涂有红色部分的长度不小于圆周总长的三分之一.

►ξ_A, ξ_B, ξ_C 分别表示 A, B, C 是红色的指标，那么 $\xi_A + \xi_B + \xi_C \geq 1$. R 表示涂红色的集合，那么三变量同分布且 $\mathsf{P}(\xi_A = 1) = \frac{|R|}{2\pi}$, 因此 $|R| \geq 2\pi/3$.

19. n 个球随机放入 m 个盒子， ξ 是空盒的个数，求 $\mathsf{E}\xi$.

►

$$\mathsf{E}(\xi_n|\xi_{n-1}) = (\xi_{n-1} - 1) \cdot \frac{\xi_{n-1}}{m} + \xi_{n-1} \cdot (1 - \frac{\xi_{n-1}}{m}) = (1 - \frac{1}{m})\xi_{n-1}.$$

因此 $\mathsf{E}\xi_n = (1 - 1/m)\mathsf{E}\xi_{n-1} = m \cdot (1 - 1/m)^n$.

20. 从 $1, 2, \cdots, n$ 中随机地选 ξ 个数，设 ξ 是 $\{1, 2, \cdots, n\}$ 上均匀分布的，求选出的数字和的期望.

X 是选出的数字之和．那么 $\mathsf{E}(X|\xi) = \frac{n+1}{2} \cdot \xi$, 因此

$$\mathsf{E}X = \frac{n+1}{2} \cdot \mathsf{E}\xi = \left(\frac{n+1}{2}\right)^2.$$

21. 设 $\xi_1, \cdots, \xi_k$ 是独立同分布的非负整数值随机变量，证明

$$\mathsf{E}\min(\xi_1, \cdots, \xi_k) = \sum_{n\geq 1}(\sum_{m\geq n} p_m)^k$$

其中 $p_m = \mathsf{P}(\xi_1 = m)$.

22. 设 ξ 是随机变量．证明： (1) 如果存在 $N > 0$ 使得 $\mathsf{E}(|\xi|; |\xi| > N) < \infty$, 则 $\mathsf{E}|\xi| < \infty$; (2) 如果 ξ 可积，则 $\lim_n \mathsf{E}(\xi; |\xi| > n) = 0$.

►(1) $\mathsf{E}|\xi| = \mathsf{E}(|\xi|; |\xi| \leq N) + \mathsf{E}(|\xi|; |\xi| > N) \leq N + \mathsf{E}(|\xi|; |\xi| > N) < \infty$.

(2) 令 $\xi_n := \xi \cdot 1_{\{|\xi|>n\}}$. 那么 $|\xi_n| \leq |\xi|$, 且 $\xi_n \xrightarrow{\text{a.s.}} 0$, 由控制收敛定理推出 $\lim_n \mathsf{E}|\xi_n| = 0$.

23. 设 ξ 是随机变量，证明： $\mathsf{E}|\xi|<\infty$ 当且仅当 $\sum_{n\geq 1}\mathsf{P}(|\xi|\geq n)<\infty$.

►因为 $n\mathsf{P}(n\leq|\xi|<n+1)\leq\mathsf{E}(|\xi|;n\leq|\xi|<n+1)\leq(n+1)\mathsf{P}(n\leq|\xi|<n+1)$, 故

$$\sum_{n\geq 0}n\mathsf{P}(n\leq|\xi|<n+1)\leq\mathsf{E}|\xi|\leq 1+\sum_{n\geq 0}n\mathsf{P}(n\leq|\xi|<n+1).$$

再仿习题 3 证明 $\sum_{n\geq 0}n\mathsf{P}(n\leq|\xi|<n+1)=\sum_{n\geq 1}\mathsf{P}(|\xi|\geq n)$.

24. 利用 Bernoulli 的大数定律证明 Weierstrass 定理：设 f 是 $[0,1]$ 上连续函数，定义 Bernstein 多项式

$$f_n(x)=\sum_{k=0}^{n}\binom{n}{k}f(k/n)x^k(1-x)^{n-k},\ x\in[0,1].$$

则 f_n 在 $[0,1]$ 上一致收敛于 f.

25. 设随机变量 ξ 的分布函数连续，期望为 μ, 中点为 m, 方差为 σ^2, 证明：$(\mu-m)^2\leq\sigma^2$.

不妨设 $m=0$. 我们要证明 $\mu^2\leq\sigma^2$, 即 $(\mathsf{E}\xi)^2\leq\mathsf{E}\xi^2-(\mathsf{E}\xi)^2$ 或者 $2(\mathsf{E}\xi)^2\leq\mathsf{E}\xi^2$. 而 $(\mathsf{E}\xi)^2=(\mathsf{E}\xi^+-\mathsf{E}\xi^-)^2$. 括号中两数均是非负的，不妨设 $\mathsf{E}\xi^+$ 更大，那么由 Cauchy-Schwarz 不等式和 0 是中点的性质，

$$(\mathsf{E}\xi)^2\leq(\mathsf{E}\xi^+)^2\leq\mathsf{E}\xi^2\cdot\mathsf{P}(\xi\geq 0)=\frac{1}{2}\mathsf{E}\xi^2.$$

26. 设 g 是 $\mathbf{R}$ 上非负递增连续函数， ξ 是随机变量. 证明：对于 $x>0$, 若 $g(x)>0$, 则有

$$\mathsf{P}(\xi>x)\leq\frac{\mathsf{E}g(\xi)}{g(x)}.$$

27. 设随机变量 ξ 是标准化的，证明：对 $x>0$,

$$\mathsf{P}(\xi\geq x)\leq\frac{1}{1+x^2}.$$

此不等式不能再改进，证明：对任何 $x>0$, 存在某概率空间上的标准化随机变量 ξ 使得

$$\mathsf{P}(\xi\geq x)=\frac{1}{1+x^2}.$$

►方法一：任取 $a>0$, 那么

$$\mathsf{E}(\xi+a)^2\geq\mathsf{E}((\xi+a)^2;\xi\geq x)\geq(a+x)^2\mathsf{P}(\xi>x),$$

因此 $\mathsf{P}(\xi>x)\leq(1+a^2)/(x+a)^2$. 给定 x, 右边的最小值是 $a=1/x$ 时达到的，等于 $1/(1+x^2)$.

►方法二 (06 曹翔宇): 令 $p = \mathsf{P}(\xi \geq x)$, 那么

$$\begin{aligned}\mathsf{E}\xi^2 &\geq \mathsf{E}(\xi^2;\xi \leq 0) + \mathsf{E}(\xi^2;\xi \geq x)\\ &\geq \mathsf{E}(\xi^2;\xi \leq 0) + x^2 p\end{aligned}$$

由 Cauchy-Schwarz 不等式,

$$\mathsf{E}(\xi^2;\xi \leq 0) \geq \frac{[\mathsf{E}(|\xi|;\xi \leq 0)]^2}{\mathsf{P}(\xi \leq 0)} \geq \frac{[\mathsf{E}(|\xi|;\xi \leq 0)]^2}{1-p},$$

而 $0 = \mathsf{E}\xi \geq \mathsf{E}(\xi;\xi \leq 0) + \mathsf{E}(\xi;\xi \geq x)$, 故 $\mathsf{E}(|\xi|;\xi \leq 0) \geq \mathsf{E}(\xi;\xi \geq x) \geq xp$, 因此

$$1 = \mathsf{E}(\xi^2) \geq \mathsf{E}(\xi^2;\xi \leq 0) + x^2 p \geq \frac{x^2p^2}{1-p} + x^2 p,$$

由此推出 $p \leq 1/(1+x^2)$.

►方法三 (06 王颖川): 令 $p = \mathsf{P}(\xi \geq x)$; 由 Chebyshev 不等式与 Cauchy 不等式,

$$\begin{aligned}x^2p^2 &\leq [\mathsf{E}(\xi;\xi \geq x)]^2 = [\mathsf{E}(\xi;\xi < x)]^2\\ &\leq \mathsf{E}(\xi^2;\xi < x)\cdot \mathsf{P}(\xi < x) = (1-\mathsf{E}(\xi^2;\xi \geq x))^2(1-p)\\ &\leq (1-x^2p)(1-p),\end{aligned}$$

因此 $p \leq 1/(1+x^2)$.

28. 一个连续函数 f 称为有紧支撑, 如果 $\{x \in \mathbf{R} : f(x) \neq 0\}$ 是有界集. 证明: 随机变量 ξ,η 有相同分布函数当且仅当对任何有紧支撑的非负连续函数 f 有 $\mathsf{E}f(\xi) = \mathsf{E}f(\eta)$.

►对任何 $a < b$, 可以构造紧支撑连续函数列 $f_n \uparrow 1_{(a,b)}$, 由单调收敛定理推出

$$\mathsf{P}(a < \xi < b) = \mathsf{E}1_{(a,b)}(\xi) = \mathsf{E}1_{(a,b)}(\eta) = \mathsf{P}(a < \eta < b).$$

推出 ξ,η 同分布.

29. 求在一个掷硬币游戏里，分别求 3-pattern (101), (111) 和 (110) 首次出现时刻的期望.

►利用条件期望性质.

$$\begin{aligned}&\mathsf{E}\xi = \mathsf{E}(\xi|0)/2 + \mathsf{E}(\xi|1)/2;\ \mathsf{E}(\xi|0) = 1 + \mathsf{E}\xi;\\ &\mathsf{E}(\xi|1) = \mathsf{E}(\xi|10)/2 + \mathsf{E}(\xi|11)/2;\ \mathsf{E}(\xi|11) = 1 + \mathsf{E}(\xi|1);\\ &\mathsf{E}(\xi|10) = \mathsf{E}(\xi|100)/2 + \mathsf{E}(\xi|101)/2;\ \mathsf{E}(\xi|100) = 3 + \mathsf{E}\xi;\ \mathsf{E}(\xi|101) = 3,\end{aligned}$$

解方程 $\mathsf{E}\xi = 10$. 其它类似. $\mathsf{E}\xi_{(111)} = 14$, $\mathsf{E}\xi_{(110)} = 8$.

30. 设 N 是取非负整数值的随机变量，$\{\xi_i\}_{i\geq 1}$ 是独立同分布可积随机变量列且与 N 独立．证明：$\mathsf{E}\sum_{i=1}^{N}\xi_i = \mathsf{E}N\cdot\mathsf{E}\xi_1$.
31. 设 N 服从参数为 λ 的 Poisson 分布，$\{\xi_n : n\geq 0\}$ 是独立同分布非负整数值随机变量也独立于 N, $S := \sum_{k=1}^{N}\xi_k$, 证明：$\mathsf{E}Sg(S) = \lambda\mathsf{E}(g(S+\xi_0)\xi_0)$.

   ►

$$\begin{aligned}
\mathsf{E}(Sg(S)) &= \sum_{k\geq 0}\mathsf{E}(Sg(S)|N=k)\mathsf{P}(N=k)\\
&= \sum_{k\geq 0}\mathsf{E}(\sum_{i=1}^{k}\xi_i g(\sum_{i=1}^{k}\xi_i))\mathsf{P}(N=k)\\
&= \sum_{k\geq 1}\mathsf{E}(\xi_0 g(\sum_{i=1}^{k-1}\xi_i+\xi_0))\frac{\lambda^k}{(k-1)!}e^{-\lambda}\\
&= \lambda\cdot\sum_{k\geq 0}\mathsf{E}(\xi_0 g(\sum_{i=1}^{k}\xi_i+\xi_0))\frac{\lambda^k}{k!}e^{-\lambda}\\
&= \lambda\cdot\mathsf{E}(\xi_0 g(S+\xi_0)).
\end{aligned}$$

32. 设随机变量 ξ 与 η 几乎处处相等，证明：(1) 它们的分布函数相等；(2) 如果 ξ 可积那么 η 也可积且 $\mathsf{E}\xi = \mathsf{E}\eta$.

第二章　随机变量与分布

§2.1　分布函数

习　　题

1. 证明： $\mathscr{B} = \sigma(\{(-\infty, x]:\ x \in \mathbf{Q}\})$.
2. 设 ξ 是 Ω 上的函数， $\mathscr{A}$ 是 $\mathbf{R}$ 的子集类，证明： $\xi^{-1}(\sigma(\mathscr{A})) = \sigma(\xi^{-1}(\mathscr{A}))$.
3. 设 $(\Omega, \mathscr{F}, \mathsf{P})$ 是概率空间， ξ 是随机变量，对任何 $B \in \mathscr{B}(\mathbf{R})$, 定义 $\mu(B) := \mathsf{P}(\xi \in B)$, 证明： $(\mathbf{R}, \mathscr{B}(\mathbf{R}), \mu)$ 也是概率空间．它称为是 ξ 的像概率空间．概率 μ 与 ξ 的分布函数是什么关系？
4. 设 ξ, η 是 Ω 上函数，如果 η 是关于 $\sigma(\xi)$ 可测的随机变量，那么存在 $\mathbf{R}$ 上 Borel 可测函数 f 使得 $\eta = f(\xi)$.

   ▶实现 $\sigma(\xi) = \xi^{-1}(\mathscr{B})$, 若 η 是示性函数， $\eta = 1_A$, 那么存在 $B \in \mathscr{B}$, 使得 $A = \{\xi \in B\}$, 因此 $\eta = 1_B(\xi)$, 成立．显然若 η 是简单的时候也成立，现在设 η 非负，那么存在简单的 η_n 使得 $\eta_n \uparrow \eta$, 且存在 Borel 函数 f_n 使得 $\eta_n = f_n(\xi)$, 令 $f = \liminf_n f_n$. 那么 f 是 Borel 函数且 $\eta = f(\xi)$.
5. 设 $(\Omega, \mathscr{F}, \mathsf{P})$ 是概率空间, 用 $\mathscr{N}$ 表示 $\mathscr{F}$ 的零概率集的子集全体, 即 $N \in \mathscr{N}$ 当且仅当存在 $N_0 \in \mathscr{F}$, 满足 $N \subset N_0$ 且 $\mathsf{P}(N_0) = 0$. 证明： $\{A \cup N:\ A \in \mathscr{F}, N \in \mathscr{N}\}$ 是 σ- 域.

   ▶令 $\mathscr{F}'$ 是要证明的集类．那么容易证明 $\emptyset, \Omega \in \mathscr{F}'$．设 $A \cup N \in \mathscr{F}'$, 即 $A \in \mathscr{F}$, $N \subset N_0 \in \mathscr{F}$ 且 $\mathsf{P}(N_0) = 0$. 那么

   $$(A \cup N)^c = (A \cup N_0)^c \cup (N_0 \setminus N),$$

   而 $A \cup N_0 \in \mathscr{F}$, $N_0 \setminus N \in \mathscr{N}$, 因此 $(A \cup N)^c \in \mathscr{F}'$. 验证 $\mathscr{F}'$ 对可列并封闭是类似的.
6. 设 ξ 是密度为 f 的连续性随机变量， ϕ 是严格递增可导函数．那么 $\phi(\xi)$ 也是连续型随机变量，密度函数是

   $$x \mapsto f(\phi^{-1}(x)) \cdot \frac{\mathrm{d}\phi^{-1}(x)}{\mathrm{d}x}.$$

   举例说明如果 ϕ 仅是递增的，那么 $\phi(\xi)$ 未必是连续型的.
7. 设 ξ 的密度函数是

   $$f(x) = \begin{cases} a + bx^2, & x \in [0,1]; \\ 0, & x \notin [0,1]. \end{cases}$$

如果 $\mathsf{E}\xi=\frac{3}{5}$, 求 a,b.

►$a=3/5$, $b=6/5$.

8. 证明：如果连续型随机变量 ξ 可积，则 (1) $\lim_{y\to+\infty} y\mathsf{P}(\xi>y)=0$; (2)

$$\mathsf{E}\xi=\int_0^\infty \mathsf{P}(\xi>y)\mathrm{d}y-\int_0^\infty \mathsf{P}(\xi<-y)\mathrm{d}y.$$

►先设 ξ 非负，验证 $\mathsf{E}\xi=\int_0^\infty \mathsf{P}(\xi>y)\mathrm{d}y$. 有两种方法 (1) 用 Fubini 定理，

$$\mathsf{E}\xi=\mathsf{E}\int_0^\xi \mathrm{d}y=\mathsf{E}\int_0^\infty 1_{\{\xi>y\}}\mathrm{d}y=\int_0^\infty \mathsf{P}(\xi>y)\mathrm{d}y.$$

(2) 如果 ξ 可积，那么用控制收敛定理推出 $\lim_{y\to\infty} y\mathsf{P}(\xi>y)=0$, 再使用分布积分公式推出结论.

►一般地， $\mathsf{E}\xi=\mathsf{E}\xi^+-\mathsf{E}\xi^-$, 而当 $y>0$ 时， $\mathsf{P}(\xi^+>y)=\mathsf{P}(\xi>y)$, $\mathsf{P}(\xi^->y)=\mathsf{P}(\xi<-y)$.

9. 证明：如果 ξ 是非负随机变量，那么

$$\mathsf{E}\xi^n=\int_0^\infty nx^{n-1}\mathsf{P}(\xi>x)\mathrm{d}x.$$

10. 如果随机变量 ξ 的分布函数 F 连续，证明： $F(\xi)$ 服从 $[0,1]$ 上均匀分布.

►对任何 $x\in(0,1)$, 定义 $F^{-1}(x):=\sup\{y: F(y)=x\}$, 因为 F 连续，故 $F^{-1}(x)\in\mathbf{R}$ 且 $F(F^{-1}(x))=x$. 因此 $\mathsf{P}(F(\xi)\le x)=\mathsf{P}(\xi\le F^{-1}(x))=F(F^{-1}(x))=x$, 均匀.

11. 设 ξ 是标准正态分布的，对 $\alpha\in(0,1)$, z_α 是标准正态分布的 α- 分位点. 证明：

$$\inf\{y-x:\ \mathsf{P}(x<\xi<y)=1-\alpha\}=2z_{\frac{\alpha}{2}}.$$

►令左边等于 z. 因为 $\mathsf{P}(-z_{\alpha/2}<\xi<z_{\alpha/2})=1-\alpha$, 故 $z\le 2z_{\alpha/2}$, 反过来，固定 $a>0$, 由于正态密度函数的性质， $\mathsf{P}(x-a/2<\xi<x+a/2)$ 当 $x=0$ 时最大. 如果 $0<y-x<2z_{\alpha/2}$, 那么

$$\mathsf{P}(x<\xi<y)\le\mathsf{P}(-\frac{y-x}{2}<\xi<\frac{y-x}{2})<\mathsf{P}(-z_{\alpha/2}<\xi<z_{\alpha/2})=1-\alpha.$$

12. 设 Φ 是标准正态分布函数. 证明：

$$(\frac{1}{x}-\frac{1}{x^3})\mathrm{e}^{-\frac{1}{2}x^2}<\sqrt{2\pi}(1-\Phi(x))<\frac{1}{x}\mathrm{e}^{-\frac{1}{2}x^2},\ x>0.$$

►导数方法.

13. 设 X 服从标准正态分布，$a>0$. 证明：$\lim_{x\to+\infty}\mathsf{P}(X>x+\frac{a}{x}|X>x)=\mathrm{e}^{-a}$.
►L'hospital
14. 设 U 服从 $[0,1]$ 上均匀分布. (1) 求 $X=[nU]+1$ 的分布. (2) 求 $X=[\frac{\log U}{\log q}]+1$ 的分布，其中 $0<q<1$.
►(1) $\{1,2,\cdots,n\}$ 上均匀分布.　(2) 几何分布.
15. 如果 X 服从参数为 λ 的 Poisson 分布，$Y\sim\Gamma(r,1)$, 那么对 $r=1,2,\ldots,$ 有 $\mathsf{P}(X\geq r)=\mathsf{P}(Y\leq\lambda)$.
►用分布积分公式计算 $\mathsf{P}(Y\leq\lambda)$.
16. 在单位圆周上固定点 A, 然后在圆周上任取点 B, (1) 求弦 AB 长度的分布与期望；(2) 求弦 AB 与 A 点切线间夹角的分布.
17. 设 X 是 Cauchy 分布，证明：$1/X$ 也是 Cauchy 分布的.
18. 设连续型随机变量的分布函数为 F, 期望为 μ. 证明：$\int_{-\infty}^{a}F(x)dx=\int_{a}^{+\infty}(1-F(x))dx$ 当且仅当 $a=\mu$.
►参考第 8 题，如果等式成立，用分部积分公式,

$$aF(a)-\int_{-\infty}^{a}x\mathrm{d}F(x)=-a(1-F(a))+\int_{a}^{\infty}x\mathrm{d}F(x),$$

因此 $a=\int_{\mathbf{R}}x\mathrm{d}F(x)=\mu$. 可以反推.
19. 设 ξ 服从参数为 $\alpha>0$ 的指数分布，求 $\eta=[\xi]$ 的分布 (其中 $[x]$ 表示 x 的整数部分).
20. 设 F 是分布函数. 证明：$\lim_{x\to+\infty}x\int_{x}^{+\infty}\frac{1}{y}\mathrm{d}F(y)=0$.
►$x\int_{x}^{\infty}y^{-1}\mathrm{d}F(y)\leq\int_{x}^{\infty}\mathrm{d}F(y)=1-F(x)$.
21. 设 F 是分布函数，$a>0$. 证明：

$$\int_{-\infty}^{+\infty}(F(x+a)-F(x))\mathrm{d}x=a.$$

►

$$\begin{aligned}\text{left side}&=\lim_{n}\int_{-n}^{n}(F(x+a)-F(x))\mathrm{d}x\\&=\lim_{n}\left(\int_{n}^{n+a}F(x)\mathrm{d}x-\int_{-n}^{-n+a}F(x)\mathrm{d}x\right)=a.\end{aligned}$$

22. 设 ξ 服从参数 α 的指数分布，计算期望 $\mathsf{E}\sin\xi$.
►$\mathsf{E}\sin\xi=\alpha/(1+\alpha^2)$.

§2.2 随机向量及联合分布

习 题

1. 设 ξ,η 是独立随机变量，证明：

$$\mathsf{P}(\xi=\eta)=\mathsf{E}\phi(\eta)=\sum_{x\in\mathbf{R}}\mathsf{P}(\xi=x)\mathsf{P}(\eta=x),$$

其中 $\phi(y):=\mathsf{P}(\xi=y)$; 因此当其中一个的分布函数连续时， $\mathsf{P}(\xi=\eta)=0$.
►也有两种方法. (1)

$$\mathsf{P}(\xi=\eta)=\lim_n\sum_{k=-\infty}^{+\infty}(F((k+1)/2^n)-F(k/2^n))(G((k+1)/2^n)-G(k/2^n)),$$

因为连续的分布函数一定是一致连续的 (请验证), 故当其中有一个连续时, $\mathsf{P}(\xi=\eta)=0$. (2) 设 $D=\{(x,x):x\in\mathbf{R}\}$, 那么由 Fubini 定理

$$\mathsf{P}(\xi=\eta)=\mathsf{E}1_D(\xi,\eta)=\int_{\mathbf{R}}\mathsf{E}1_D(\xi,x)\mathrm{d}F_\eta(x)=\sum_x\mathsf{P}(\xi=x)\mathsf{P}(\eta=x).$$

2. (1) 设随机变量 ξ_1,ξ_2,ξ_3,ξ_4 独立同分布，分布函数连续，求 ξ_1,ξ_2 为端点的开区间与 ξ_3,ξ_4 为端点的开区间不相交的概率. (2) 设 X_1,X_2,Y_1,Y_2 是独立随机变量， X_1,X_2 同分布，分布函数与密度函数分别是 F 与 f, Y_1,Y_2 同分布，分布函数与密度函数分别是 G 与 g, 证明：区间 (X_1,X_2) 与 (Y_1,Y_2) 不相交的概率为

$$\frac{1}{3}+\mathsf{E}\left[(F(X_1)-G(X_1))^2+(F(Y_1)-G(Y_1))^2\right].$$

►(1) 由上题，四个点中任何两点都不可能重合，因此四个点的排列方式有 $4!=24$ 种，所求概率为 $8/24$.

3. 设 X,Y 是独立同分布非负随机变量，分布函数是 F, 如果对任何 $a,b>0$, $\inf(aX,bY)$ 与 $\frac{ab}{a+b}X$ 同分布，证明： F 是指数分布函数.

4. 设 (X,Y) 是 2 维随机向量，如果对任何 2 阶正交矩阵 Q, $(X,Y)Q$ 与 (X,Y) 有相同的分布， (或说 (X,Y) 有正交不变的分布,) 且 $\mathsf{P}(X=0,Y=0)=0$, 证明： $\mathsf{P}(Y=0)=1$. 这样随机变量 $\frac{X}{Y}$ 与 $\frac{X}{|Y|}$ 有意义，证明它们有相同分布并求它们共同的分布.
►(1) 令 A 是除去 0 点的 x 轴. 那么对任何平面上旋转 α 角的变换 Q_α 有 $Q_\alpha A\cap Q_\beta A=\emptyset$, $0<\alpha<\beta<2\pi$. 由旋转不变性

$$1=\mathsf{P}((X,Y)\in\mathbf{R}^2)\geq\mathsf{P}((X,Y)\in\bigcup_n Q_{1/n}A)$$

$$= \sum_n \mathsf{P}((X,Y) \in Q_{1/n}A) = \sum_n \mathsf{P}((X,Y) \in A),$$

因此 $\mathsf{P}((X,Y) \in A) = 0$, 而 $\mathsf{P}(Y=0) = \mathsf{P}((X,Y) \in A) + \mathsf{P}(X=0, Y=0) = 0$.
►(2) 对任何 $z \in \mathbf{R}$,

$$\mathsf{P}(\frac{X}{Y} \le z) = \mathsf{P}(X \le zY; Y > 0) + \mathsf{P}(X \ge zY; Y < 0);$$
$$\mathsf{P}(\frac{X}{|Y|} \le z) = \mathsf{P}(X \le zY; Y > 0) + \mathsf{P}(X \le -zY; Y < 0).$$

由正交不变性, (X,Y) 与 $(-X,Y)$ 同分布, 因此 $\mathsf{P}(X \ge zY; Y < 0) = \mathsf{P}(-X \ge zY; Y < 0) = \mathsf{P}(X \le -zY; Y < 0)$.
►(3) 用旋转不变性证明 $\arctan(Y/X)$ 是 $(-\pi/2, \pi/2)$ 上均匀分布的, 因此 Y/X 是 Cauchy 分布的 (参考例 2.1.7). 当 $\alpha < \beta$ 时, 用 $S(\alpha,\beta)$ 表示幅角从 α 到 β 的扇形. 旋转不变性蕴含着对任何 $\theta \in \mathbf{R}$,

$$\mathsf{P}((X,Y) \in S(\alpha,\beta)) = \mathsf{P}((X,Y) \in S(\alpha+\theta, \beta+\theta)).$$

不难推出可加性: 只要 $\alpha_i \ge 0$, $\alpha_1 + \alpha_2 \le 2\pi$, 则

$$\mathsf{P}((X,Y) \in S(0, \alpha_1+\alpha_2)) = \mathsf{P}((X,Y) \in S(0,\alpha_1) + \mathsf{P}((X,Y) \in S(0,\alpha_2)).$$

这蕴含着对 $x \ge 0$, 且 $x(\beta-\alpha) \le 2\pi$,

$$\mathsf{P}((X,Y) \in S(x\alpha, x\beta)) = x \cdot \mathsf{P}((X,Y) \in S(\alpha,\beta)).$$

因此当 $\beta - \alpha \le 2\pi$ 时,

$$\begin{aligned}\mathsf{P}((X,Y) \in S(\alpha,\beta)) &= \mathsf{P}((X,Y) \in S(0, \beta-\alpha)) \\ &= \frac{\beta-\alpha}{2\pi}\mathsf{P}((X,Y) \in S(0,2\pi)) = \frac{\beta-\alpha}{2\pi}.\end{aligned}$$

那么对 $\alpha \in (-\pi/2, \pi/2)$, $\arctan(Y/X) \le \alpha$ 当且仅当

$$(X,Y) \in S(-\pi/2, \alpha) \cup S(\pi/2, \alpha+\pi).$$

因此

$$\mathsf{P}(\arctan(Y/X) \le \alpha) = 2\mathsf{P}((X,Y) \in S(-\pi/2,\alpha)) = \frac{2 \cdot (\alpha + \pi/2)}{2\pi} = \frac{\pi/2 + \alpha}{\pi}.$$

5. (1) 证明：

$$B(r_1, r_2) = \frac{\Gamma(r_1)\Gamma(r_2)}{\Gamma(r_1 + r_2)}.$$

(2) 求参数为 r_1, r_2 的 β 分布的随机变量的期望和方差.

6. 设 X, Y 是独立且具有相同指数分布的随机变量，求 $\frac{Y}{X}$ 的密度函数.

►$\mathsf{P}(Y/X \le z) = \mathsf{P}(Y \le zX) = \int_{x>0} \mathrm{d}x \int_0^{zx} \alpha^2 \mathrm{e}^{-\alpha(x+y)} \mathrm{d}y = \frac{z}{1+z}$.

7. 设 (X, Y) 是上单位半圆上均匀分布的，求 $Z = \frac{Y}{X}$ 的分布.

8. 设 (X, Y) 服从区域 D 上均匀分布，其中 D 是点 $(0,0), (1,0), (1,1)$ 所围成的三角形．求 $\rho(X, Y)$.

9. 在 $[0,1]$ 上随机取点 ξ 把它分成为两段，记 $U := \xi \wedge (1 - \xi)$, $V = 1 - U$, 分别是短段与长段的长度．求 $\frac{V}{U}$ 的密度.

►密度： $2/(1+x)^2$, $x \ge 1$.

10. 在一个线段 AB 上任取三点 X, Y, Z, 求线段 AX, AY, AZ 能够组成三角形的概率． (2) 在一个线段上任取两点自然形成三个线段，求它们能组成三角形的概率.

►(1) $\{(x, y, z) : x, y, z \in [0,1], x + y \ge z, y + z \ge x, z + x \ge y\}$ 的体积为 $1/2$.
(2) A 是组成三角形的事件． $\mathsf{P}(A) = \mathsf{P}(A; \xi < \eta) + \mathsf{P}(A; \eta < \xi)$. 现在

$$\begin{aligned}
&\mathsf{P}(A, \xi < \eta) \\
&\quad = \mathsf{P}(\xi < \eta, \xi + (\eta - \xi) > 1 - \eta, \xi + (1 - \eta) > \eta - \xi, (1 - \eta) + (\eta - \xi) > \xi) \\
&\quad = \mathsf{P}(\xi < \eta, \eta > 1/2, \xi < 1/2, \eta - \xi < 1/2) = 1/8.
\end{aligned}$$

因此 $\mathsf{P}(A) = 1/4$.

11. 设 $\xi = (\xi_1, \xi_2, \xi_3)$ 是在 $\mathbf{R}^3$ 的单位球面上均匀分布的，求 ξ_1 的边缘分布.

►Case $n = 3$.

$$\mathsf{P}(\xi_3 \le x) = \frac{1}{4\pi} \int_{x_1^2 + x_2^2 + x_3^2 = 1} dS = \frac{x+1}{2},$$

应用极坐标的方法计算面积分．因此 ξ_3 在 $[-1, 1]$ 上均匀分布.

12. 设 $\mathbf{X} = (X_1, \cdots, X_n)$ 是独立的参数为 p 的 Bernoulli 随机变量．设 $f, g : \{0,1\}^n \to \mathbf{R}$ 是递增的． (1) 令 $e(p) := \mathsf{E} f(\mathbf{X})$, 证明： e 关于 p 也是递增的． (2) 证明： $\mathrm{cov}(f(\mathbf{X}), g(\mathbf{X})) \ge 0$.

►(1) 设 $p_1 < p_2$, $Y_1, \cdots, Y_n$ 是参数 p_2 的 Bernoulli, $Z_1, \cdots, Z_n$ 是参数 p_1/p_2 的 Bernoulli 且两个 Bernoulli 独立，令 $Y_i' = Y_i Z_i$, 那么 $Y_1', \cdots, Y_n'$ 是参数 p_1 的 Bernoulli 且 $Y \ge Y'$. 因此 $e(p_1) = \mathsf{E} f(Y') \le \mathsf{E} f(Y) = e(p_2)$.

►(2) 应用归纳法和下面的计算

$$\begin{aligned}\operatorname{cov}(f(X),g(X)) &= \mathsf{E}f(X)g(X)-\mathsf{E}f(X)\mathsf{E}g(X)\\ &= q\mathsf{E}f(X',0)g(X',0)+p\mathsf{E}f(X',1)g(X',1)\\ &\quad -(q\mathsf{E}f(X',0)+p\mathsf{E}f(X',1))(q\mathsf{E}g(X',0)+p\mathsf{E}g(X',1))\\ &= q\cdot\operatorname{cov}(f(X',0),g(X',0))+p\cdot\operatorname{cov}(f(X',1),g(X',1))\\ &\quad +pq\cdot(\mathsf{E}f(X',1)-\mathsf{E}f(X',0))(\mathsf{E}g(X',1)-\mathsf{E}g(X',0)).\end{aligned}$$

13. 设 X,Y 是随机变量，期望 0, 方差 1, 协方差 ρ. 证明：

$$\mathsf{E}(\max\{X^2,Y^2\})\le 1+\sqrt{1-\rho^2}.$$

►应用 Cauchy-Schwarz 不等式

$$\begin{aligned}\mathsf{E}\max(X^2,Y^2) &= \frac{1}{2}\mathsf{E}(X^2+Y^2+|X^2-Y^2|)\\ &= 1+\frac{1}{2}\mathsf{E}|X-Y||X+Y|\\ &\le 1+\frac{1}{2}\left(\mathsf{E}(X-Y)^2\mathsf{E}(X+Y)^2\right)^{\frac{1}{2}}\\ &= 1+\sqrt{1-\rho^2}.\end{aligned}$$

14. 设 $\mathbf{X}$ 是 n- 维随机 (行) 向量，$V(\mathbf{X})$ 是其协方差矩阵, 证明： $\det(V(\mathbf{X}))=0$ 当且仅当存在非零的 $\mathbf{a}\in\mathbf{R}^n$, $b\in\mathbf{R}$ 使得 $\mathsf{P}(\mathbf{a}\mathbf{X}^{\mathrm{T}}=b)=1$.

►不妨设 $\mathsf{E}X=0$. 如果存在非零的 $a\in\mathbf{R}^n$, 使得 $a^TX=0$, 那么 $a^TXX^Ta=0$, 或者 $a^T(\mathsf{E}XX^T)a=0$, 因此矩阵 $\mathsf{E}XX^T$ 非满秩的. 反之，如果 $\mathsf{E}XX^T$ 奇异，那么存在非零的向量 a 使得 $a^T\mathsf{E}XX^Ta=0$. 因此 $\mathsf{E}|a^TX|^2=0$, 即 $a^TX=0$ a.s.

15. 设 ξ,η 独立, 分别服从参数 a,b 的 Cauchy 分布, 证明 $\xi+\eta$ 服从参数为 $a+b$ 的 Cauchy 分布. （提示：利用复变的留数定理计算积分.) 因此 Cauchy 分布有再生性.

►$\xi+\eta$ 的密度函数是

$$\frac{ab}{\pi^2}\int_{\mathbf{R}}\frac{1}{(x-y)^2+a^2}\cdot\frac{1}{y^2+b^2}\mathrm{d}y.$$

应用留数的方法计算积分.

16. 设 $\mathbf{X}=(\xi_1,\cdots,\xi_n)$ 是服从 n- 维正态分布的随机向量，Q 是秩为 m 的 $n\times m$-阶矩阵，证明： $\mathbf{X}Q$ 是 m- 维正态分布的随机向量，也就是说，正态分布随机向量的非退化线性变换仍然是正态分布的.

17. 设随机变量 (ξ,η) 独立且都有连续的密度, 如果 (ξ,η) 得分布是旋转不变的, 证明：它们服从正态分布.
18. 设随机变量 ξ,η 独立，$\xi\sim\chi^2(n),\eta\sim\chi^2(m)$, 求随机变量 $\frac{\eta/m}{\xi/n}$ 的密度函数. 这个分布称为是参数为 m,n 的 F 分布.
19. 设 $\{\xi_n : n\geq 1\}$ 是独立同分布随机序列，且分布函数是连续的，用 N 表示使得 $\xi_{n+1}>\xi_1$ 成立的最小指标 n, $N:=\inf\{n:\xi_{n+1}>\xi_1\}$, 求 N 的分布.
 ▶对任何 n, $N>n$ 当且仅当对所有 $1\leq i\leq n$, $\xi_{i+1}<\xi_1$. 因此类似习题 2 的理由推出
$$\mathsf{P}(N>n)=\mathsf{P}(\xi_{i+1}>\xi_1,1\leq i\leq n)=\frac{n!}{(n+1)!}=1/(n+1).$$
20. (参考例 ??) 在单位圆内均匀随机地取个点，求以此点为中点的弦所割的短圆弧的密度函数.
 ▶取点 P, P 到圆心的距离为 r, 弦长为 $2\sqrt{1-r^2}$. 设 θ 是它所对的圆心角, $0<\theta<\pi$, 那么由余弦定理 $4(1-r^2)=2-2\cos\theta$, 因此
$$\begin{aligned}\mathsf{P}(\theta\leq x)&=\mathsf{P}(\cos\theta\geq\cos x)\\&=\mathsf{P}(2r^2-1\geq\cos x)\\&=\mathsf{P}(r^2\geq\frac{1+\cos x}{2})\\&=1-\frac{1}{2}(1+\cos x)=\frac{1}{2}(1-\cos x).\end{aligned}$$
21. 设 $(X_1,\cdots,X_n)$ 是独立标准正态分布的. Ψ 是向量 $(X_1,\cdots,X_n)$ 与一个固定向量的夹角，求 Ψ 的密度函数.
 ▶固定向量 $\mathbf{e}_1=(1,0,\cdots,0)$. 那么 $\cos\Psi=\frac{X_1}{|X|}$. 我们来算 $\cos\Psi$ 的密度，让 $x\in(-1,1)$,
$$\mathsf{P}(\frac{|X_1|}{|X|}\leq x)=\int\cdots\int_{|y_1|\leq x|y|,y\in\mathbf{R}^n}(2\pi)^{-n/2}\exp\{-\frac{1}{2}\}|y|^2\mathrm{d}y.$$
 而积分区域等于 $|y_1|\leq\sqrt{y_2^2+\cdots+y_n^2}\cdot\frac{x}{\sqrt{1-x^2}}$. 上面对 x 求导得密度函数为
$$\begin{aligned}&2(2\pi)^{-n/2}\cdot(1-x^2)^{-3/2}\cdot\int \mathrm{e}^{-\frac{1}{2}(y_2^2+\cdots+y_n^2)}\cdot(y_2^2+\cdots+y_n^2)^{1/2}\mathrm{d}y_2\cdots\mathrm{d}y_n\\&=C_n\cdot(1-x^2)^{(n-3)/2},\end{aligned}$$
 其中 $1/C_n=\int_0^1(1-x^2)^{(n-3)/2}\mathrm{d}x$. 然后 $\frac{X_1}{|X|}$ 的密度容易得到. 当 $n=3$ 时, 正好是均匀分布.

22. 设 (X,Y) 服从 2 维正态分布， $\mathsf{E}X=\mathsf{E}Y=0$, $\mathsf{E}X^2=\mathsf{E}Y^2=1$, $\rho(X,Y)=\rho$.

(a) 证明： X 与 $\frac{Y-\rho X}{\sqrt{1-\rho^2}}$ 是独立的标准正态分布的.

(b) 令 ϕ 是 (X,Y) 的幅角，证明 ϕ 的密度函数为

$$\frac{\sqrt{1-\rho^2}}{2\pi(1-2\rho\sin\theta\cos\theta)},\ 0\le\theta<2\pi.$$

(c) 再求 $\mathsf{P}(X>0,Y>0)$.

(d) 令 $Z=\max(X,Y)$, 证明： $\mathsf{E}Z=\sqrt{(1-\rho)\pi^{-1}}$, $\mathsf{E}Z^2=1$.

►(1) 首先二维正态分布的线性变换仍然是二维正态分布的, 因此 (X,Y_1) (其中 $Y_1=\frac{Y-\rho X}{\sqrt{1-\rho^2}}$) 仍然是二维正态分布的，然后

$$\mathrm{D}Y_1=\mathsf{E}Y_1^2=\frac{1}{1-\rho^2}\mathsf{E}Y^2-2\rho\mathsf{E}XY+\rho^2\mathsf{E}X^2=1$$

且 $\mathrm{cov}(X,Y_1)=0$.

►(2) 令 $\tan\theta_0=\frac{-\rho}{\sqrt{1-\rho^2}}$,

$$\begin{aligned}\mathsf{P}(X>0,Y>0)&=\int_{x>0,y>0}\frac{1}{2\pi\sqrt{1-\rho^2}}\cdot\exp\left\{-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right\}\mathrm{d}x\mathrm{d}y\\&=\int_{x>0,\sqrt{1-\rho^2}y+\rho x>0}\frac{1}{2\pi}\exp\{-\frac{1}{2}(x^2+y^2)\}\mathrm{d}x\mathrm{d}y\\&=\int_{\theta_0}^{\pi/2}\mathrm{d}\theta\int_0^\infty\frac{1}{2\pi}r\mathrm{e}^{-\frac{r^2}{2}}\mathrm{d}r\\&=\frac{1}{2\pi}(\pi/2-\theta_0)=\frac{1}{2\pi}(\pi/2+\arcsin\rho),\end{aligned}$$

其中作积分变换 $(x,y)\mapsto(x,\frac{y-\rho x}{\sqrt{1-\rho^2}})$.

►(3) 由对称性且作如 (2) 的积分变换

$$\begin{aligned}\mathsf{E}Z&=2\int_{y>x}yf(x,y)\mathrm{d}x\mathrm{d}y\\&=\frac{1}{\pi}\int_{\sqrt{1-\rho^2}y>-\rho x+x}(\sqrt{1-\rho^2}y+\rho x)\mathrm{e}^{-\frac{1}{2}(x^2+y^2)}\mathrm{d}x\mathrm{d}y\\&=\frac{1}{\pi}\sqrt{1-\rho^2}\cdot I(\sqrt{\frac{1-\rho}{1+\rho}})+\frac{1}{\pi}\rho\cdot(-I(\sqrt{\frac{1+\rho}{1-\rho}}))\\&=\sqrt{(1-\rho)/\pi}.\end{aligned}$$

利用恒等式 (验证): 当 $\alpha>0$ 时,

$$I(\alpha)=\int_{y>\alpha x}y\mathrm{e}^{-\frac{1}{2}(x^2+y^2)}\mathrm{d}x\mathrm{d}y=\sqrt{\frac{2\pi}{1+\alpha^2}}.$$

类似地

$$\mathsf{E}Z^2 = \frac{1}{\pi}\int_{y>\frac{1-\rho}{sqrt1-\rho^2}x}(\sqrt{1-\rho^2}y+\rho x)^2\cdot \mathrm{e}^{-\frac{1}{2}(x^2+y^2)}\mathrm{d}x\mathrm{d}y,$$

令

$$J(\alpha) := \frac{1}{2\pi}\int_{y>\alpha x} y^2\mathrm{e}^{-\frac{1}{2}(x^2+y^2)}\mathrm{d}x\mathrm{d}y,\ \alpha>0.$$

那么 $J'(\alpha) = -\frac{1}{2\pi}\int_{\mathbf{R}}(\alpha x)^2\mathrm{e}^{-\frac{1}{2}(1+\alpha^2)x^2}x\mathrm{d}x = 0$, 因此 $J(\alpha)$ 与 α 无关，$J(\alpha) = J(0) = 1/2$. 而且不难验证

$$\int_{y>\alpha x} xy\mathrm{e}^{-\frac{1}{2}(x^2+y^2)}\mathrm{d}x\mathrm{d}y = 0,$$

因此

$$\mathsf{E}Z^2 = \frac{1}{\pi}(2\pi(1-\rho^2)\cdot\frac{1}{2}+2\pi\rho^2\cdot\frac{1}{2}) = 1.$$

23. (1) 在一个圆心为 O 的单位圆周上任取两点 A,B, 求形成的三角形 OAB 面积的分布与期望.　(2) 在单位圆周上任取三点 A,B,C, 求三角形 ABC 面积的期望.

►(1) θ 服从 $(-\pi,\pi)$ 上均匀分布，$S_{OAB} = |\sin\theta|$, 其密度为 $\frac{2}{\pi}\frac{1}{\sqrt{1-x^2}}$, 期望为 $2/\pi$.

►(2) 把 A 固定为坐标原点. 圆心在 $(1,0)$ 点.　θ_1,θ_2 分别是 y 轴与 AB, AC 的夹角，他们独立且服从 $(0,\pi)$ 上均匀分布. 那么

$$S_{ABC} = 2|\sin\theta_1\cdot\sin\theta_2\cdot\sin(\theta_2-\theta_1)|.$$

然后计算 $\mathsf{E}S_{ABC} = 3/2\pi$.

24. 设随机变量 $\xi_1,\xi_2,\cdots,\xi_{m+n}$ $(n>m)$ 是独立的，有相同的分布并且有有限的方差. 试求 $S=\xi_1+\xi_2+\cdots+\xi_n$ 与 $T=\xi_{m+1}+\xi_{m+2}+\cdots+\xi_{m+n}$ 两和之间的相关系数.

►$\rho(S,T) = 1-m/n$.

25. 若随机变量 ξ 的密度函数是偶函数，且 $\mathsf{E}\xi^2<\infty$. 试证 $|\xi|$ 与 ξ 不相关，但它们不相互独立.

26. 设 $\xi_1,\cdots,\xi_n$ 是独立同分布随机变量，分布函数 F 连续. 证明：顺序统计量的第 i 个 $\xi_{(i)}$ 的分布函数是

$$F_i^*(x) = \frac{n!}{(i-1)!(n-i)!}\int_0^{F(x)} t^{i-1}(1-t)^{n-i}\mathrm{d}t.$$

如果 F 是 $(0,1)$ 上均匀分布的，那么 $\xi_{(i)}$ 服从参数为 $i, n-i+1$ 的 β 分布.
提示：证明对任何 $x \in (0,1)$ 和 $0 < i \le n$, 有

$$\frac{n!}{(i-1)!(n-i)!}\int_0^x t^{i-1}(1-t)^{n-i}\mathrm{d}t = \sum_{k=i}^n \binom{n}{k} x^k (1-x)^{n-k}.$$

►简单的分布积分计算得

$$\binom{n}{i}\int_0^x i\cdot t^{i-1}\cdot(1-t)^{n-i}\mathrm{d}t = \binom{n}{i}x^i(1-x)^{n-i} + \binom{n}{i+1}\int_0^x (i+1)t^i(1-t)^{n-i-1}\mathrm{d}t,$$

然后再用归纳法.

27. 设 $\xi_1, \cdots, \xi_n$ 是独立同分布随机变量，分布函数 F 连续.
    (a) 求顺序统计量 $\xi_{(1)}, \cdots, \xi_{(n)}$ 的联合分布密度;
    (b) 如果 ξ_i 服从参数为 α 的指数分布，证明：$\{\xi_{(i)} - \xi_{(i-1)} : 1 \le i \le n\}$ (其中 $\xi_{(0)} := 0$) 独立，并求它们的分布.
    (c) 设 F 是均匀分布且 $i < j$, 证明：$\xi_{(i)}/\xi_{(j)}$ 与 $\xi_{(j)}$ 独立.

►(1) 联合密度

$$n! f(x_1)\cdots f(x_n)\cdot 1_{\{x_1<x_2<\cdots<x_n\}}.$$

►(2) 取 $x_i \ge 0$,

$$\begin{aligned}
&\mathsf{P}(\xi_{(i)} - \xi_{(i-1)} > x_i, 1 \le i \le n) \\
&\quad = n! \int\cdots\int_{y_i - y_{i-1} > x_i, 1\le i\le n} e^{-y_1-\cdots-y_n} dy_1\cdots dy_n \\
&\quad = n! e^{-x_n} \int\cdots\int_{y_i - y_{i-1} > x_i, 1\le i\le n-1} e^{-y_1-\cdots-y_{n-2}-2y_{n-1}} dy_1\cdots dy_{n-1} \\
&\quad = \cdots = \exp\left\{-\sum_{i=1}^n (n-i+1)x_n\right\}.
\end{aligned}$$

►(3) $(\xi_{(k)})$ 的联合密度为 $n! 1_{\{0<x_1<\cdots<x_n<1\}}$. 对除 $\xi_{(i)}, \xi_{(j)}$ 外的所有变量积分得这两者得联合密度为

$$n!\cdot\frac{x_i^{i-1}}{(i-1)!}\cdot\frac{(x_j-x_i)^{j-i-1}}{(j-i-1)!}\cdot\frac{(1-x_j)^{n-j}}{(n-j)!}.$$

对于 $x, y \in (0,1)$,

$$\mathsf{P}(\xi_{(i)}/\xi_{(j)} \le x, \xi_{(j)} \le y) = C\cdot\iint_{x_i \le x x_j, x_j \le y} x_i^{i-1}(x_j - x_i)^{j-i-1}(1-x_j)^{n-j} dx_i dx_j$$

先对 x 求导再对 y 求导得 $\xi_{(i)}/\xi_{(j)}$ 与 $\xi_{(j)}$ 得联合密度为

$$C\cdot x^{i-1}(1-x)^{j-i-1}\cdot y^{j-1}(1-y)^{n-j},$$

因此两者独立且都是 beta 分布的.

28. 设 $X=(X_1,\cdots,X_n)$ 是 n 维平方可积的随机向量，期望为零，协方差矩阵 Σ 是非退化的．证明：若 S 是 $\mathbf{R}^n$ 的一个紧凸集，那么

$$\mathsf{P}(X\in S)\leq\frac{1}{1+\inf\{x\Sigma^{-1}x^{\mathrm{T}}:x\in S\}}.$$

(提示：先考虑一维的情况．)

29. 在三角形 $\triangle_1$ 中随机取三点连成三角形 $\triangle_2$, 证明： $\mathsf{E}|\triangle_2|=|\triangle_1|/12$.

►(05 朱世衡) 也就是证明 $\mathsf{E}(|\triangle_2|/|\triangle_1|)=1/12$. 这个值在仿射变换下不变．现在建立直角坐标系，取点 $O(0,0)$, $M(0,1)$, $N(1,0)$ 组成三角形 $\triangle_1$.

设随机放入的前三个点按横坐标从小到大依次为 $A(a,a'),B(b,b'),C(c,c')$, 用 S 表示 $\triangle ABC$ 的面积，现在要算 S 的期望 $\mathsf{E}S$.

先把 S 用 a,b,c,a',b',c' 表示出来．延长 AB 交直线 $x=c$ 于 $C_0(c,c_0')$, 那么 S 等于三角形 ACC_0 的面积减去 BCC_0 的面积，即

$$S=\frac{1}{2}(c-a)\cdot|c'-c_0'|-\frac{1}{2}(c-b)\cdot|c'-c_0'|=\frac{1}{2}(b-a)\cdot|c'-c_0'|.$$

我们要计算积分

$$\mathsf{E}S=8\int_0^1da\int_a^1db\int_b^1dc\int_0^{1-a}da'\int_0^{1-b}db'\int_0^{1-c}dc'(b-a)\cdot|c'-c_0'|.$$

关于 $c'-c_0'$ 的符号，有四种情况： (1) C_0 在 ON 下方， (2) C_0 在 MN 上方； C_0 在 $\triangle OMN$ 内部，但 (3) $c'>c_0'$, 或者 (4) $c'<C_0'$. 看 AB 线的方程为

$$AB:\ y-a'=\frac{b'-a'}{b-a}(x-a),$$

这时

$$c_0'=a'+\frac{b'-a'}{b-a}(c-a).$$

如果 $c_0'=0$, 那么 $b'=\frac{c-b}{c-a}a'$, 记其为 b_1. 如果 $c_0'=1-c$, 那么

$$b'=\frac{c-b}{c-a}a'+(1-c)\cdot\frac{b-a}{c-a},$$

记其为 b_2. 先对 b', c' 积分，我们先算下面的积分

$$\begin{aligned}
&\int_0^{1-b}\int_0^{1-c}|c'-c_0'|db'dc' \\
&= \int_0^{b_1}\int_0^{1-c}(c'-c_0') + \int_{b_1}^{b_2}\int_0^{c_0'}(c_0'-c') \\
&\quad + \int_{b_1}^{b_2}\int_{c_0'}^{1-c}(c'-c_0') + \int_{b_2}^{1-b}\int_0^{1-c}(c_0'-c')db'dc' \\
&= .....
\end{aligned}$$

最后的答案 1/72 是用 Mathematica 算出来的. 考虑三个点的顺序共有 6 种.

30. 设 ξ, η 独立，证明：如果 $\xi+\eta$ 可积，那么 ξ, η 都可积.
►设 $f(y) := \mathsf{E}|\xi+y|$, 由 Fubini 定理， $\mathsf{E}f(\eta) = \mathsf{E}|\xi+\eta|$, 因此 $f(\eta)$ 可积，至少存在一个 y, 使得 $f(y) < \infty$, 从而 $\mathsf{E}|\xi| \le \mathsf{E}|\xi+y| + |y| < \infty$.

§2.3 条件期望与独立性

习 题

1. 设 ξ, η 是独立同分布的可积的离散随机变量. 证明： $\mathsf{E}(\xi|\xi+\eta) = \frac{1}{2}(\xi+\eta)$.
►利用定义证明 $\mathsf{E}(\xi|\xi+\eta) = \mathsf{E}(\eta|\xi+\eta)$.
2. 设 $\mathscr{A}$ 是子 σ- 域， $A \in \mathscr{A}$, 证明： $\mathsf{E}(1_A\xi|\mathscr{A}) = 1_A\mathsf{E}(\xi|\mathscr{A})$.
►先设 ξ 是示性函数，简单随机变量，然后用单调收敛定理.
3. 证明： (1) $(\mathsf{E}(\xi|\mathscr{A}))^2 \le \mathsf{E}(\xi^2|\mathscr{A})$. (2) $\mathsf{E}(\mathsf{E}(\xi|\mathscr{A}))^2 \le \mathsf{E}\xi^2$.
►证明恒等式：

$$\mathsf{E}((\xi-\mathsf{E}(\xi|\mathscr{A}))^2|\mathscr{A}) = \mathsf{E}(\xi^2|\mathscr{A}) - (\mathscr{E}(\xi|\mathscr{A}))^2.$$

然后 (1), (2) 都是显然的.

4. 设 X, Y 独立的且服从标准正态分布的， $Z = X+Y$. 求 $\mathsf{E}(Z|X>0, Y>0)$.
►首先 $\mathsf{P}(X>0, Y>0) = 1/4$, 然后

$$\mathsf{P}(X+Y \le z, X>0, Y>0) = \frac{1}{2\pi}\int_0^{\infty} \mathrm{e}^{-\frac{y^2}{2}}\mathrm{d}y \int_0^{z-y} \mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x,$$

求导得密度函数 (但得不到具体的解析表达式). 期望

$$\mathsf{E}(X+Y|X>0, Y>0) = 4\int_{x>0}\int y>0(x+y)\frac{1}{2\pi}\mathrm{e}^{-\frac{x^2+y^2}{2}}\mathrm{d}x\mathrm{d}y = 4/\sqrt{2\pi}.$$

5. 设 (X,Y) 服从区域 D 上均匀分布，其中 D 是点 $(0,0),(1,0),(1,1)$ 所围成的三角形．求条件期望 $\mathsf{E}(Y|X)$.
6. 设 ξ 服从 $N(a,\sigma^2)$, Φ 是标准正态分布函数，求 $\mathsf{E}\Phi(\xi)$. (提示：取独立于 ξ 且服从标准正态的随机变量 U, 证明： $\Phi(\xi)=\mathsf{P}(U\le\xi|\xi)$.)
7. 设 (ξ,η) 有联合密度函数，证明：

$$\mathsf{E}(\xi|\eta)=\int_{\mathbf{R}} sf_{\xi|\eta}(s|\eta)\mathrm{d}s.$$

8. 设 (X,Y) 服从二维正态分布，$\mathsf{E}X=\mathsf{E}Y=0, \mathsf{E}X^2=\sigma^2, \mathsf{E}Y^2=\tau^2, \rho(X,Y)=\rho$. 证明：
   (a) $\mathsf{E}(X|Y)=\dfrac{\rho\sigma}{\tau}Y$;
   (b) $\mathsf{E}(X|X+Y)=\dfrac{\sigma^2+\rho\sigma\tau}{\sigma^2+2\rho\sigma\tau+\tau^2}(X+Y)$.

   ►(1) 不妨假设 X,Y 是标准化的．那么

$$f_{X|Y}(x|y)=f(x,y)/f(y)=\frac{1}{\sqrt{2\pi(1-\rho^2)}}\exp\{-\frac{(x-\rho y)^2}{2(1-\rho^2)}\}.$$

   即 $X|Y=y$ 服从 $N(\rho y,1-\rho^2)$ 分布，因此 $\mathsf{E}(X|Y)=\rho Y$. (2) 利用 (1).

$$\mathsf{E}(\frac{X}{\sigma}|X+Y)=\rho(X,X+Y)\cdot\frac{X+Y}{\sqrt{X+Y}}.$$

9. 设 X 服从参数为 n, U 的二项分布，其中 U 服从 $[0,1]$ 上均匀分布．求 X 的分布．

   ►对 $0\le k\le n$,

$$\begin{aligned}\mathsf{P}(X=k)&=\mathsf{E}(\mathsf{P}(X=k|U))=\mathsf{E}\binom{n}{k}U^k(1-U)^{n-k}\\&=\int_0^1\binom{n}{k}x^k(1-x)^{n-k}\mathrm{d}x\\&=\binom{n}{k}B(k+1,n-k+1)=\frac{1}{n+1}.\end{aligned}$$

10. 设连续型可积随机变量 X 的密度函数为 f, 写出 $\mathsf{E}(X||X|)$.

    ►对 $x>0$,

$$\begin{aligned}\mathsf{E}(X||X|=x)&=\lim_{\delta\to0}\mathsf{E}(X||X|\in(x-\delta,x+\delta))\\&=\lim_{\delta\to0}\frac{\int_{|y|\in(x-\delta,x+\delta)}yf(y)\mathrm{d}y}{\int_{|y|\in(x-\delta,x+\delta)}f(y)\mathrm{d}y}\end{aligned}$$

$$= \frac{xf(x) - xf(-x)}{f(x) + f(-x)},$$

因此

$$\mathsf{E}(X||X|) = \frac{|X|f(|X|) - |X|f(-|X|)}{f(|X|) + f(-|X|)}.$$

另外一个方法，存在函数 g 使得 $\mathsf{E}(X||X|) = g(|X|)$, 用定义算 g.

11. 设 $\{Y_k : 1 \le k \le n\}$ 是独立且标准正态分布的， $X_i := \sum_{j=1}^n c_{ij}Y_j$, 其中 $\{c_{ij}\}$ 是常数．证明

$$\mathsf{E}(X_j|X_k) = \left(\frac{\sum_i c_{ji}c_{ki}}{\sum_i c_{ki}^2}\right) X_k.$$

12. 设 X, Y, Z 服从 3 维正态分布, 边缘分布都是标准正态的， $\rho_1 = \rho(X, Y), \rho_2 = \rho(Y, Z), \rho_3 = \rho(Z, X)$. (1) 证明：

$$\mathsf{P}(X > 0, Y > 0, Z > 0) = \frac{1}{8} + \frac{1}{4\pi}\sum_{i=1}^{3} \arcsin \rho_i.$$

(2) 设 $\rho_1 = \rho(X, Y)$, 证明：

$$\mathsf{E}(Z|X, Y) = \frac{(\rho_3 - \rho_1\rho_2)X + (\rho_2 - \rho_1\rho_3)Y}{1 - \rho_1^2}.$$

►(1) 利用 §2.2 的习题 22 计算, 首先 (X, Y, Z) 与 $-(X, Y, Z)$ 有相同的分布, 那么

$$\begin{aligned}
\mathsf{P}(X > 0, Y > 0, Y > 0) &= 1 - \mathsf{P}(X < 0) - \mathsf{P}(Y < 0) - \mathsf{P}(Z < 0) \\
&\quad + \mathsf{P}(X < 0, Y < 0) + \mathsf{P}(Y < 0, Z < 0) + \mathsf{P}(Z < 0, X < 0) \\
&\quad - \mathsf{P}(X < 0, Y < 0, Z < 0) \\
&= 1 - 2/3 + \frac{1}{2\pi}\sum_{i=1}^{3} \arcsin \rho_i - \mathsf{P}(X > 0, Y > 0, Z > 0),
\end{aligned}$$

故而得证． (2) 直接算协方差矩阵证明

$$Z - \frac{1}{1 - \rho_1^2}[(\rho_3 - \rho_1\rho_2)X + (\rho_2 - \rho_3\rho_1)Y]$$

与 X, Y 独立，因而推出结论.

13. (1) 设 ξ 是可积随机变量， $\mathscr{A}$ 是 $\mathscr{F}$ 的子 σ- 域，证明： $|\mathsf{E}(\xi|\mathscr{A})| \le \mathsf{E}(|\xi||\mathscr{A})$. (2) 设 ξ, η 可积，独立且期望为零，证明： $\mathsf{E}|\xi + \eta| \ge \mathsf{E}|\xi|$. (提示：利用 (1).)

14. 设有随机变量集 $\{\xi_i : i \in I\}$, 如果
$$\mathscr{A} := \bigcup_{I_0} \sigma(\xi_i : i \in I_0),$$
其中 I_0 取遍 I 的所有有限子集, 证明：事件类 $\mathscr{A}$ 对交封闭且 $\sigma(\xi_i : i \in I) = \sigma(\mathscr{A})$.
15. 一根木棍随机断为两段，然后较长的一根又随机地断为两段，求此三段能组成三角形的概率.
►设 ξ 是第一个断点，它是 $[0.1]$ 上均匀分布的， $X = \xi \vee (1-\xi)$ 是长的一段． Y 是 $[0, X]$ 上均匀分布的．那么 (X, Y) 的联合密度为
$$f(x,y) = f_X(x) f_{Y|X}(y|x) = 2 \cdot 1_{[1/2,1]}(x) \cdot \frac{1}{x} 1_{[0,x]}(y).$$
三线段 $1-X, Y, X-Y$ 能够组成三角形当且仅当 $1-X+Y > X-Y$, $1-X+X-Y > Y$, $Y+(X-Y) > 1-X$, 即 $X-Y < 1/2$, $Y < 1/2$. 概率为
$$\mathsf{P}(X-Y<1/2, Y<1/2) = \int_{y<x, 1/2<x<1} 1_{\{x-y<1/2, y<1/2\}} 2/x dx dy = \ln 4 - 1.$$
16. (1) 设两个非负随机序列 $\{\xi_n : n \geq 1\}$ 与 $\{\eta_n : n \geq 1\}$ 独立, 证明： $\liminf_n \xi_n$ 与 $\liminf_n \eta_n$ 独立． (2) 证明定理 **??** 中 $\{\eta_n : n \geq 1\}$ 是独立随机序列
►因为两个随机序列独立，由定理 2.3.4 后的说明， $\mathscr{F}_1 = \sigma(\xi_n : n \geq 1)$ 与 $\mathscr{F}_2 = \sigma(\eta_n : n \geq 1)$ 独立，而 $\liminf_n \xi_n, \liminf \eta_n$, 分别关于 $\mathscr{F}_1$ 与 $\mathscr{F}_2$ 可测，因此他们独立.
17. 随机序列 $\{\xi_n : n \geq 1\}$ 的尾 σ- 域定义为
$$\bigcap_{n \geq 1} \sigma(\{\xi_k : k \geq n\}).$$
一个 σ- 域服从 0-1 律，如果其中的事件的概率非零即一．证明：独立随机序列的尾 σ- 域服从 0-1 律.
18. 设 ξ, η 是有界随机变量，用 Weierstrass 逼近定理证明 ξ, η 独立当且仅当对任何非负整数 m, n 有 $\mathsf{E}\xi^m\eta^n = \mathsf{E}\xi^m \cdot \mathsf{E}\eta^n$.
19. 设 ξ, η 是独立随机变量．如果 ξ 与 $\xi - \eta$ 也独立，证明 ξ 等于常数.
20. 设 η 是随机变量, 且 $\{\xi, \eta\}$ 与 $\mathscr{A}$ 独立. 则 $\mathsf{E}(\xi|\eta, \mathscr{A}) = \mathsf{E}(\xi|\eta)$, 其中 $\mathsf{E}(\xi|\eta, \mathscr{A})$ 表示 ξ 关于 $\sigma(\sigma(\eta) \cup \mathscr{A})$ 的条件期望.
21. 设 Z 是在半径为 a 的圆内随机取两点的距离，求 $\mathsf{E}Z$, $\mathsf{E}Z^2$.

►先设两点为 (X_1, Y_1) 与 (X_2, Y_2), 他们服从单位圆上均匀分布. 先算容易的 $\mathsf{E}Z^2$,

$$\begin{aligned}\mathsf{E}Z^2 &= \mathsf{E}[(X_1 - X_2)^2 + (Y_1 - Y_2)^2] \\ &= 2\mathsf{E}(X_1 - X_2)^2 = 4\mathsf{E}X_1^2.\end{aligned}$$

而

$$\mathsf{E}X_1^2 = \frac{1}{2}\mathsf{E}(X_1^2 + Y_1^2) = \frac{1}{2\pi}\int\int_{x^2+y^2\le 1}(x^2+y^2)\mathrm{d}x\mathrm{d}y = \frac{1}{4}.$$

►$\mathsf{E}Z$ 要困难得多，我们用 Crafton 的方法.

►(1) 在半径为 a 的圆内任取一点 P 到圆周上给定点 A 的距离的期望 $\mathsf{E}|AP|$. 把圆心放在 $(a, 0)$ 处，使用极坐标表示 $r \le 2a\cos\theta$, $\theta \in [-\pi/2, \pi/2]$. 其均匀分布密度为 $r\mathrm{d}r\mathrm{d}\theta/\pi a^2$. 因此

$$\mathsf{E}|AP| = \frac{1}{\pi a^2}\int_{-\pi/2}^{\pi/2}\int_0^{2a\cos\theta} r\cdot r\mathrm{d}r\mathrm{d}\theta = \frac{32a}{9\pi}.$$

►(2) 设 D_x 是半径为 x 的圆， $A = D_{x+h}\setminus D_x$. 取两点记为 X, Y, 令

$$\lambda(x) := \mathsf{E}(Z|X\in D_x, Y\in D_x).$$

那么

$$\begin{aligned}&\lambda(x+h) = \mathsf{E}(Z|X\in D_{x+h}, Y\in D_{x+h}) \\ &= \frac{\mathsf{E}(Z; X\in D_{x+h}, Y\in D_{x+h})}{\mathsf{P}(X\in D_{x+h}, Y\in D_{x+h})} \\ &= \frac{\mathsf{E}(Z; X\in D_x, Y\in D_x) + 2\mathsf{E}(Z; X\in D_x, Y\in A) + o(h)}{\mathsf{P}(X\in D_{x+h}, Y\in D_{x+h})} \\ &= \lambda(x)\cdot\frac{\mathsf{P}(X\in D_{x+h}, Y\in D_{x+h})}{\mathsf{P}(X\in D_{x+h}, Y\in D_{x+h})} \\ &+ 2\mathsf{E}(Z|X\in D_x, Y\in A)\cdot\frac{\mathsf{P}(X\in D_x, Y\in A)}{\mathsf{P}(X\in D_{x+h}, Y\in D_{x+h})} + o(h) \\ &= \lambda(x)\cdot\frac{x^4}{(x+h)^4} + 2\mathsf{P}(X\in D_x, Y\in A)\cdot\frac{x^2((x+h)^2-x^2)}{(x+h)^4} + o(h) \\ &= \lambda(x)\cdot(1 - \frac{4h}{(x+h)} + o(h)) + 2\mathsf{P}(X\in D_x, Y\in A)\cdot\frac{x^2(2hx+h^2)}{(x+h)^4} + o(h).\end{aligned}$$

利用 (1) 的结果得

$$\lambda'(x) = \lim_{h\to 0}\frac{\lambda(x+h)-\lambda(x)}{h} = \lambda(x)\cdot(-4/\lambda) + 2\cdot\frac{32x}{9\pi}\cdot\frac{2}{x}.$$

解方程，初值 $\lambda(0) = 0$, 得 $\lambda(x) = \frac{128x}{45\pi}$.

22. 设 S 是一个半径为 a 的圆内随机取三个点组成的三角形的面积. 证明

$$\mathsf{E}S = \frac{35a^2}{48\pi}.$$

►(应该放在 §2.3) 应用 Crafton 的方法. 先设一个点 P 在圆周上. 还是应用极坐标计算 $\mathsf{E}|PQR|$, 答案是 $\frac{35a^2}{36\pi}$. 然后类似 16 题建立一个带初值的微分方程

$$\begin{cases}\lambda'(x) &= -\frac{6}{x}\lambda(x) + \frac{35x}{6\pi};\\ \lambda(0) &= 0.\end{cases}$$

§2.4 随机变量的收敛

习 题

1. 设 $\{A_n\}$ 是独立事件列, $\xi(\omega) := |\{n : \omega \in A_n\}|$, 即包含有 ω 的集合 A_n 的个数. 证明: $\mathsf{P}(\xi < \infty) > 0$ 蕴含着 $\mathsf{E}\xi < \infty$.
   ►$\xi = \sum_{n\geq 1} 1_{A_n}$. $\mathsf{E}\xi = \sum_{n\geq 1} \mathsf{P}(A_n)$, 再利用 Borel-Cantelli 引理.
2. 设 $\{A_n\}$ 是事件列, 证明: $\mathsf{P}(\limsup_n A_n) \geq \limsup \mathsf{P}(A_n)$.
   ►应用 Fatou 引理
3. 设 $\{\xi_n\}$ 是随机序列, 满足 $\sum_{n\geq 1} \mathsf{E}|\xi_n| < \infty$. 证明: $\sum_{n\geq 1} \xi_n$ 几乎处处收敛于一个可积函数 ξ, 且 $\mathsf{E}\xi = \sum_{n\geq 1} \mathsf{E}\xi_n$.
   ►令 $\eta = \sum_{n\geq 1} |\xi_n|$, 那么 $\mathsf{E}\eta < \infty$, 因此 $\mathsf{P}(\eta < \infty) = 1$, 即 $\sum_{n\geq 1} \xi_n$ 几乎处处绝对收敛, 然后再应用控制收敛定理.
4. 设 $\xi_n \xrightarrow{\mathrm{p}} \xi$ 且 $\xi_n \xrightarrow{\mathrm{p}} \eta$, 证明: $\mathsf{P}(\xi = \eta) = 1$.
5. 设 $\{\xi_n\}$ 是独立同分布随机序列, 服从 $(0, 1)$ 上均匀分布, 令 Y_n 是 $\xi_1, \xi_2, \cdots, \xi_{2n+1}$ 按大小排列后位于中间的那个, 证明: Y_n 几乎处处收敛于 1/2.
   ►对很小的 $\epsilon > 0$, 估计 $\mathsf{P}(|Y_n - 1/2| > \epsilon)$. 由 §2.2 的 22 题, Y_n 服从参数为 $n+1, n+1$ 的 β- 分布, 应用 Stirling 公式,

$$\begin{aligned}\mathsf{P}(|Y_n - 1/2| > \epsilon) &= 2 \cdot \frac{(2n+1)!}{(n!)^2} \int_0^{1/2-\epsilon} t^n (1-t)^n \mathrm{d}t \\ &\leq 2 \cdot (2n+1) \cdot 2^{2n+1} \cdot \frac{1}{\sqrt{n}} \cdot (1/2-\epsilon)^n (1/2+\epsilon)^n \cdot (1/2-\epsilon) \\ &= 2 \cdot (2n+1) \cdot \sqrt{2/n} \cdot (1/2-\epsilon)(1-(2\epsilon)^2)^n,\end{aligned}$$

   因此 $\sum_n \mathsf{P}(|Y_n - 1/2| > \epsilon)$ 收敛, 推论 2.4.2 说明, Y_n 几乎处处收敛于 1/2.

6. (1) 证明：如果存在极限为零的正数列 $\{\delta_n\}$ 使得 $\sum_n \mathsf{P}(|\xi_n-\xi|>\delta_n)<\infty$，那么 $\xi_n \xrightarrow{\text{a.s.}} \xi$. (2) 设 $\{\xi_n\}$ 是独立同分布随机序列，服从参数 1 的指数分布，令 $Y_n=\inf\{\frac{\xi_k}{k}:1\le k\le n\}$，计算 $\mathsf{P}(Y_n>n^{-3/2})$ 并证明 $\sum_{n\ge1} Y_n$ 几乎处处收敛.

►

$$\begin{aligned}\mathsf{P}(Y_n>n^{-3/2}) &= \mathsf{P}(\xi_k/k>n^{-3/2}, 1\le k\le n)\\ &= \prod_{k=1}^{n}\mathsf{P}(\xi_k>k\cdot n^{-3/2}) = \prod_{k=1}^{n}\mathrm{e}^{-k\cdot n^{-3/2}}\\ &= \mathrm{e}^{-n+1/2\sqrt{n}},\end{aligned}$$

因此 $\sum_n \mathsf{P}(Y_n>n^{3/2})<\infty$，由 Borel-Cantelli 引理，对几乎所有的 $\omega, Y_n(\omega)>n^{-3/2}$ 最多只有有限多个成立，因此 $\sum_n Y_n$ 几乎处处收敛.

7. 设 $\{A_n\}$ 是事件列， m 是固定自然数．用 G 表示至少属于 m 个 A_n 中的元素全体．证明： G 可测且 $m\mathsf{P}(G)\le\sum_n \mathsf{P}(A_n)$.

►令 $\xi=\sum 1_{A_n}$，那么 $G=\{\xi\ge m\}$，应用 Chebyshev 不等式.

8. 设 $\{\xi_n\}$ 是随机变量列，被一个可积随机变量控制或者满足 $\sup_n \mathsf{E}|\xi_n|^2<\infty$，证明： $\{\xi_n\}$ 是一致可积的.

►应用 Cauchy-Schwarz 不等式和 Chebyshev 不等式，

$$\mathsf{E}(|\xi_n|;|\xi_n|>N)\le\sqrt{\mathsf{E}|\xi_n|^2\cdot\mathsf{P}(|\xi_n|>N)}\le\mathsf{E}|\xi_n|^2\cdot\frac{1}{N}.$$

9. (1) (控制收敛定理) 如果 $\{\xi_n\}$ 被可积的 η 控制且 $\xi_n\xrightarrow{\text{p}}\xi$，证明： $\xi_n\xrightarrow{L^1}\xi$. (2) 设 ξ_n,ξ 是非负可积随机变量， $\xi_n\xrightarrow{\text{p}}\xi$, $\mathsf{E}\xi_n\longrightarrow\mathsf{E}\xi$. 证明： $\xi_n\xrightarrow{L^1}\xi$.

►(1) 首先，如果 $\{\xi_n\}$ 被可积的 η 控制且 $\xi_n\xrightarrow{\text{p}}\xi$，那么 $\{\xi_n\}$ 是一致可积的，故 $\xi_n\xrightarrow{L^1}\xi$.

(2) 显然 $(\xi-\xi_n)^+\le\xi$ 且 $(\xi-\xi_n)^+\xrightarrow{\text{p}}0$，因此 $(\xi-\xi_n)^+\xrightarrow{L^1}0$.

(3) 由 (2) 和条件， $\mathsf{E}|\xi-\xi_n|^-=\mathsf{E}[(\xi-\xi_n)^+-(\xi-\xi_n)]=\mathsf{E}(\xi-\xi_n)^+-(\mathsf{E}\xi-\mathsf{E}\xi_n)$ 趋向于 0.

10. 设 $\xi_n\xrightarrow{\text{p}}\xi$, g 是连续函数，证明： $g(\xi_n)\xrightarrow{\text{p}}g(\xi)$.

►给定 $\epsilon>0$，存在 $N>0$ 使得 $\mathsf{P}(|\xi|>N)<\epsilon$. 而 g 在 $[-N-1,N+1]$ 上绝对连续，存在 $\delta>0$ 使得只要 x,y 在这个区间内且 $|x-y|<\delta$，则 $|g(x)-g(y)|<\epsilon$. 现在只要 n 充分大使得 $\mathsf{P}(|\xi_n-\xi|>\delta)<\epsilon$ 就可以了.

►注：换成依分布收敛也是对的.

11. 设 ξ_n 是非负递减随机变量序列， $\xi_n\xrightarrow{\text{p}}0$. 证明： $\xi_n\xrightarrow{\text{a.s.}}0$.

12. 重复两个结果成功 S 失败 F 的 Bernoulli 试验，给定的一个系列，如 SFS, 在试验中无限次出现的概率是多少？

    ►设 A_n 是 $3n+1, 3n+2, 3n+3$ 三次试验的结果是 SFS. 那么 $\{A_n\}$ 独立且 $\mathsf{P}(A_n) = p^2q > 0$. 因此 $\sum \mathsf{P}(A_n) = \infty$, 由 Borel-Cantelli 引理， A_n 会无穷多次出现.

13. 证明：如果 $\{F_n\}$ 弱收敛，则其极限唯一.

14. 设分布函数列 $\{F_n\}$ 弱收敛于连续的分布函数 F. 证明：收敛对于 $x \in \mathbf{R}$ 是一致的.

    ►首先，连续的分布函数必定一致连续. 对任何 $\epsilon > 0$, 存在 $N > 0$ 和 $\delta > 0$ 使得 $F(-N) < \epsilon$, $1 - F(N) < \epsilon$ 且只要 $x < y$ 是 $[-N, N]$ 中距离小于 δ 的两点，则 $F(y) - F(x) < \epsilon$. 现在取 $-N = x_0 < x_1 < \cdots < x_{u-1} < x_u = N$ 使得连续两点的距离小于 δ. 然后 n 充分大 (只与 ϵ 有关), 使得对所有 i, $F_n(x_i) - F(x_i)| < \epsilon$. 现在可以证明对所有 x, $|F_n(x) - F(x)| < 5\epsilon$. 当 $x < -N$ 时 ($x > N$ 类似),

$$|F_n(x) - F(x)| \leq F_n(x) + F_n(x) \leq F_n(-N) + F(-N) \leq \epsilon + 2F(-N) < 3\epsilon,$$

    当 $x \in [-N, N]$, 那么存在 i 使得 $x_i \leq x < x_{i+1}$,

$$\begin{aligned}|F_n(x) - F(x)| &\leq F_n(x) - F_n(x_i) + |F_n(x_i) - F(x_i)| + F(x) - F(x_i) \\ &\leq 2\epsilon + F_n(x_{i+1}) - F_n(x_i) \\ &\leq 2\epsilon + |F_n(x_{i+1}) - F(x_{i+1})| + F(x_{i+1}) - F(x_i) + |F(x_i) - F_n(x_i)| \\ &\leq 5\epsilon,\end{aligned}$$

15. 举例说明随机变量列依分布收敛未必依概率收敛. 若 $\{\xi_n\}$ 在同一个概率空间上且 a 是常数，证明： $\xi_n \Rightarrow a$ 蕴含着 $\xi_n \xrightarrow{\mathrm{p}} a$.

16. 分布函数列 F_n 弱收敛当且仅当存在 $\mathbf{R}$ 的一个稠子集 D 使得对任何 $x \in D$, $F_n(x)$ 收敛.

    ►设 F_n 在 D 上收敛，记 $\hat{F}(x) := \lim F_n(x)$, $x \in D$, 它递增. 用 F 表示 $\hat{F}$ 的右连续化，即

$$F(x) := \inf\{\hat{F}(y) : y \in D, y > x\}, \ x \in \mathbf{R}.$$

    事实上, 对任意 $x \in \mathbf{R}$, 取 $x_n \downarrow x$, 对任意 $x < y \in D$, 当 n 充分大, 有 $x_n < y$, 故 $F(x_n) \leq \hat{F}(y)$, 即 $\lim_n F(x_n) \leq \hat{F}(y)$, 推出 $F(x) \leq \lim_n F(x_n) \leq F(x)$.

现在我们证明 F_n 弱收敛于 F. 给定 $x \in \mathbf{R}$, 取 $x_1 < x < x_2$, $x_1, x_2 \in D$, 那么 $F_n(x_1) \le F_n(x) \le F_n(x_2)$, 故

$$\hat{F}(x_1) = \lim F_n(x_1) \le \liminf F_n(x) \le \limsup F_n(x) \le \lim F_n(x_2) = \hat{F}(x_2).$$

如果 F 在 x 点连续,对任何 $\epsilon > 0$, 一定存在 $x_1' < x < x_2'$ 使得 $F(x_2') - F(x_1') < \epsilon$. 因为 D 稠密，可以做到 $x_1' < x_1 < x_2 < x_2'$, 因此

$$\hat{F}(x_2) - \hat{F}(x_1) \le F(x_2') - F(x_1') < \epsilon.$$

由此推出 $\lim_n F_n(x)$ 存在等于 $F(x)$.

另外一个方法看起来更简单些 (由 07 一位同学提供): 令

$$F_0(x) := \inf\{\hat{F}(y) : y \in D, y \ge x\},\ x \in \mathbf{R}.$$

那么 F_0 在 D 上等于 $\hat{F}$. 然后证明 F_n 在 F_0 的连续点上趋于 F_0. 再记 F_0 的右连续化为 F, 那么 F_0 与 F 在 F_0 的连续点上相等, 即 F_n 在一个稠密集上收敛于递增右连续函数 F, 从而弱收敛.

17. 满足 $\lim_{|x|\to\infty} f(x) = 0$ 的连续函数全体记为 $C_\infty(\mathbf{R})$, 它包含紧支撑连续函数空间. 证明： (Helly-Bray 定理的推广) 分布函数列 F_n 弱收敛于递增右连续函数 F 蕴含着对任何 $f \in C_\infty(\mathbf{R})$, $\int f \mathrm{d}F_n$ 收敛于 $\int f \mathrm{d}F$. 举例说明逆命题不对.

►设 $F_n \xrightarrow{\mathrm{w}} F$. 对任何自然数 N, 令

$$F_n^{(N)} := 1_{[N,\infty)} + F_n \cdot 1_{[-N,N)},$$
$$F^{(N)} := 1_{[N,\infty)} + F \cdot 1_{[-N,N)}.$$

那么 $F_n^{(N)} \xrightarrow{\mathrm{w}} F^{(N)}$. 因两边都是分布函数，对 $f \in C_\infty(\mathbf{R})$, 定理 2.4.9 推出

$$\int f \mathrm{d}F_n^{(N)} \longrightarrow \int f \mathrm{d}F^{(N)}.$$

给定 $\epsilon > 0$, 存在 N, 使得当 $|x| > N$ 时， $|f(x)| < \epsilon$. 而

$$\left|\int f \mathrm{d}F_n^{(N)} - \int f \mathrm{d}F_n\right| \le f(-N)F_n(-N) + f(N)(1 - F_n(N)) + \int_{|x|\ge N} |f| \mathrm{d}F_n \le 2\epsilon.$$

因此

$$\left|\int f \mathrm{d}F_n - \int f \mathrm{d}F\right| \le 4\epsilon + \left|\int f \mathrm{d}F_n^{(N)} - \int f \mathrm{d}F^{(N)}\right|.$$

18. 分布函数列 $\{F_n\}$ 称为是在无穷远处等度连续的，如果对任何 $\epsilon > 0$, 存在 $N > 0$ 使得对任何 $n \geq 1$, 有 $F_n(N) - F_n(-N) > 1 - \epsilon$. 设 F_n 弱收敛于递增右连续函数 F, 证明： F 是分布函数当且仅当 $\{F_n\}$ 在无穷远处等度连续.
19. 设 F 是 $[0,1]$ 上均匀分布函数. 对 $[0,1]$ 上密度趋于零的分划列 $\{(0 = x_0^{(n)} < \cdots < x_{j_n}^{(n)} = 1) : n \geq 1\}$, 任取 $a_i^{(n)} \in (x_{i-1}^{(n)}, x_i^{(n)}]$, $i = 1, \cdots, j_n$, 定义

$$F_n = \sum_{i=1}^{j_n} (x_i^{(n)} - x_{i-1}^{(n)}) 1_{[a_i^{(n)}, +\infty)}.$$

证明： $F_n \xrightarrow{\mathrm{w}} F$.

►

$$\begin{aligned} F_n &= \sum_{i=1}^{j_n} (x_i^{(n)} - x_{i-1}^{(n)}) 1_{[a_i^{(n)}, \infty)} \\ &= \sum_{i=1}^{j_n} x_i^{(n)} 1_{[a_i^{(n)}, \infty)} - \sum_{i=0}^{j_n - 1} x_i^{(n)} 1_{[a_{i+1}^{(n)}, \infty)} \\ &= \sum_{i=1}^{j_n - 1} x_i^{(n)} 1_{[a_i^{(n)}, a_{i+1}^{(n)})} + 1_{[a_{j_n}^{(n)}, \infty)}. \end{aligned}$$

设 $x \in [0,1]$, 当分划长度趋向于零时，

$$\begin{aligned} |F_n(x) - x| &\leq \sum_{i=1}^{j_n - 1} |x_i^{(n)} - x| \cdot 1_{[a_i^{(n)}, a_{i+1}^{(n)})}(x) + x 1_{[0, a_1^{(n)})}(x) + (1-x) 1_{[a_{j_n}^{(n)}, \infty)} \\ &\leq a_1^{(n)} + (1 - a_{j_n}^{(n)}) + \max_{1 \leq i \leq j_n} |x_i^{(n)} - x_{i-1}^{(n)}| \longrightarrow 0. \end{aligned}$$

20. 设 F_n 是分布函数， F 是绝对连续的分布函数且 $F_n \xrightarrow{\mathrm{w}} F$, f 是 $\mathbf{R}$ 上有界且几乎处处连续的函数. 证明： $\int f \mathrm{d}F_n \longrightarrow \int f \mathrm{d}F$, 试由此推出 $[0,1]$ 上几乎处处连续的有界 Borel 可测函数是 Riemann 可积的. (此题限于学过 Lebesgue 测度的读者.)

►设 F 是 ξ 的分布函数，那么 F 绝对连续蕴含着对任何 $\mathbf{R}$ 的零测集 N 有 $\mathsf{P}(\xi \in N) = 0$. 现在 $F_n \xrightarrow{\mathrm{w}} F$, 由 Skorohod 定理，存在 ξ_n, ξ 使得其分布函数分别是 F_n, F 且 ξ_n 点点收敛于 ξ. 存在设 f 是几乎处处连续的有界函数， $C(f)$ 是连续点集，因为 $C(f)$ 是零测集的余集，故由控制收敛定理

$$\mathsf{E} f(\xi_n) = \mathsf{E}(f(\xi_n); \xi \in C(f)) \longrightarrow \mathsf{E} f(\xi).$$

这蕴含着 $\int f \mathrm{d}F_n \longrightarrow \int f \mathrm{d}F$, 即 Riemann 和收敛.

21. 证明：分布函数列 $\{F_n\}$ 弱收敛当且仅当其所有弱收敛子列都弱收敛于同一个函数 F.

§2.5 母函数与特征函数

习 题

1. 设 $\{u_n\}$ 的母函数是 $U(t)$, 计算数列 $\{2u_n\}$, $\{u_n+1\}$, $\{nu_n\}$, $\{u_{2n}\}$ 的母函数.

   ►$G_{2\{u_n\}}(t)=2U(t)$, $G_{\{u_n+1\}}=U(t)+\frac{1}{1-t}$
2. 写出由母函数 $U(t)=\sqrt{1-4pqt^2}$ $(p,q>0,p+q=1)$ 决定的数列表达式.
3. 取 n 个相互独立的事件 A_i, 其中 $p_i=\mathsf{P}(A_i)$, $1\le i\le n$. 用 ξ 表示其中发生的事件数，求 ξ 的母函数与期望.
4. 求 Pascal 分布的母函数.
5. (Coupon 问题) 某种方便面使用有卡通人物的卡片进行推销，一套共 N 张，在每包方便面中随机地放置一张卡通卡片. 用 ξ 表示收全一套卡片所需买的方便面包数，求 ξ 的母函数.
6. 求例 ?? 中 $\{p_n\}$ 的极限.
7. 证明：如果随机变量 ξ 的特征函数的模恒等于 1, 那么 ξ 是常数.

   ►设 η 与 ξ 独立同分布，那么 $\xi-\eta$ 的特征函数恒等于 1, 故 $\xi-\eta\equiv 0$. 那么 ξ 与自身独立，即为常数.
8. 用母函数方法证明公式：

$$\sum_{j+k=l}\binom{m+j-1}{j}\binom{n+k-1}{k}=\binom{m+n+l-1}{l}.$$

   ►只需证明

$$\sum_{j\ge 0}\binom{m+j-1}{j}t^j=(1-t)^{-m}.$$

9. 连续地掷一个正面概率为 p 的硬币， X 是到 HTH 出现为止所掷的次数, Y 是到 HTH 或者 THT 出现为止所掷的次数. 证明： $\mathsf{E}s^X=\frac{p^2qs^3}{1-s+pqs^2-pq^2s^3}$ 并求 $\mathsf{E}s^Y$.

   ►

$$\mathsf{E}t^X=\mathsf{E}(t^X|H)p+\mathsf{P}(t^X|T)q=\mathsf{E}(t^X|H)p+\mathsf{P}(t^{X+1})q;$$
$$\mathsf{E}(t^X|H)=\mathsf{E}(t^X|HH)p+\mathsf{E}(t^X|HT)q=\mathsf{E}(t^{X+1}|H)p+\mathsf{E}(t^X|HT)q;$$

$$\mathsf{E}(t^X|HT) = \mathsf{E}(t^X|HTH)p + \mathsf{E}(T^X|HTT)q = t^3p + \mathsf{E}t^{X+3}q,$$

解方程得

$$\mathsf{E}t^X = \frac{qp^2t^3}{1-t+t^2pq-pq^2t^3}.$$

同理得

$$\mathsf{E}t^Y = \frac{t^3pq-t^5p^2q^2}{1-t+t^2pq-t^4p^2q^2}.$$

10. 设 $\{\xi_n : n \geq 1\}$ 是独立且都服从参数为 λ 的指数分布，对于给定的 $t>0$, 令 $\eta := \sup\{n : S_n \leq t\}$, 其中 $S_0 = 0$, $S_n = \sum_{i=1}^n \xi_i$. 证明 η 服从参数为 λt 的 Poisson 分布.
11. 一个随机变量 X 称为是格分布的，如果存在 a 与 $b>0$ 使得 X 支撑在格 $\{a+nb : n = \cdots, -1, 0, 1, 2, \cdots\}$ 上. 设 X 的特征函数为 ϕ. 证明：
   (a) X 是格分布的当且仅当存在 $x \neq 0$ 使得 $|\phi(x)| = 1$;
   (b) 如果存在不可公度的 x, x' (即 $x \neq 0, x' \neq 0, x/x'$ 是无理数) 使得 $|\phi(x)| = |\phi(x')| = 1$, 则 X 是常数.

   ►(1) 设存在 $t_0 \neq 0$, 使得 $|\phi(t_0)| = 1$. 先设 $\phi(t_0) = 1$, 那么 $\mathsf{E}\cos t_0 X = 1$, 任取 $h \in (0,1)$,

$$\mathsf{E}\cos t_0 X \leq h\mathsf{P}(\cos t_0 X \leq h) + \mathsf{P}(\cos t_0 X > h),$$

   推出 $\mathsf{P}(\cos t_0 X \leq h) = 0$, 因此 $\mathsf{P}(\cos t_0 X < 1) = 0$ 或者 $\mathsf{P}(\cos t_0 X = 1) = 1$, 即 X 分布在 $\{x : \cos t_0 x = 1\}$ 上, 格分布的. 一般地, 存在 θ_0 使得 $\mathrm{e}^{i\theta_0}\phi(t_0) = 1$, 即 $\mathsf{E}\mathrm{e}^{i(\theta_0/t_0 + X)t_0} = 1$, 因此 $X + \theta_0/t_0$ 是格分布的， X 也是格分布的.

   ►(2) 现在由 (1), $|\phi_X(x)| = 1$ 蕴含着存在 θ 使得 $xX + \theta$ 分布在 $2\pi\mathbf{Z}$; $|\phi_X(x')| = 1$ 蕴含着存在 θ' 使得 $x'X + \theta'$ 分布在 $2\pi\mathbf{Z}$. 反证法，如果 X 不是常数，那么存在 $k_1, k_2 \in \mathbf{Z}$ 使得 X 以正概率分布在 $(2k_1\pi - \theta)/x$ 与 $(2k_2\pi - \theta)/x$ 上. 自然地存在 $k_1', k_2' \in \mathbf{Z}$ 使得

$$\frac{2k_1\pi - \theta}{x} = \frac{2k_1'\pi - \theta'}{x'}, \ \frac{2k_2\pi - \theta}{x} = \frac{2k_2'\pi - \theta'}{x'}.$$

   因此 $(k_1 - k_2)/x = (k_1' - k_2')/x'$, 与不可公度的条件矛盾.
12. 设 ϕ 是连续型随机变量 X 的特征函数，证明： 对任何 $t \in \mathbf{R}$,

$$\mathrm{Re}(1-\phi(t)) \geq \frac{1}{4}\mathrm{Re}(1-\phi(2t)).$$

13. 设 ϕ 是一个特征函数，证明： $\overline{\phi}, \phi^2, |\phi|^2, \mathrm{Re}(\phi)$ 都是特征函数，而 $|\phi|$ 不一定是特征函数.

►设 X 地特征函数是 ϕ, 那么 $-X$ 的特征函数是 $\bar{\phi}$, 取与 X 独立同分布随机变量 Y, $X+Y$ 的特征函数 ϕ^2, $X-Y$ 的特征函数是 $|\phi|^2$, 设 X 的分布函数是 F, $-X$ 的分布函数是 G, 那么 $(F+G)/2$ 的特征函数是 $\mathrm{Re}\phi$. 但是 $|\phi|$ 未必是特征函数. 例如设 X 均匀分布在 $\{-1,1\}$ 上，那么 $\phi(x)=\cos x$, 但 $|\cos x|$ 不是特征函数，反设是，那么有 Y 使得 $|\cos x| = \mathsf{E}\mathrm{e}^{ixY}$, 因为左边在 0 点二次可导，因此 Y 平方可积，那么其特征函数必然处处二次可导，但 $|\cos x|$ 不是.

14. 设 ϕ 是一个有密度函数 f 的分布函数的特征函数，且设 f 是紧支撑连续函数，

 (a) 证明： $\lim_{x\to\infty}\phi(x)=0$;

 (b) 证明： (Parseval 等式) 如果 $\int_{\mathbf{R}} f(x)^2\mathrm{d}x<\infty$, 那么

$$\frac{1}{2\pi}\int_{\mathbf{R}}|\phi(t)|^2\mathrm{d}t=\int_{\mathbf{R}}f(x)^2\mathrm{d}x.$$

►设 ϕ 是密度 f 得特征函数. 先设 $f\in C_c^\infty(\mathbf{R})$, 这时分部积分,

$$\phi(x)=\int \mathrm{e}^{ixy}f(y)\mathrm{d}y=\frac{1}{ix}\int e^{ixy}f'(y)\mathrm{d}y,$$

因此 $|\phi(x)|\le\frac{1}{|x|}\|f'\|_{L^1}\longrightarrow 0$. 对一般 f, 和任何 $\epsilon>0$, 存在 $g\in C_c^\infty(\mathbf{R})$ 使得 $\|f-g\|_{L^1}<\epsilon$. 然后

$$|\phi(x)|\le\|f-g\|+|\int \mathrm{e}^{ixy}g(y)\mathrm{d}y|$$

极限是零.

►应用 Fubini 定理

$$\begin{aligned}\int\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2\sigma^2}}|\phi(t)|^2\mathrm{d}t &= \int_{\mathbf{R}^2}f(x)f(y)\mathrm{d}x\mathrm{d}y\int_{\mathbf{R}}\int\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{t^2}{2\sigma^2}}\mathrm{e}^{it(x-y)}\mathrm{d}t\\ &=\int_{\mathbf{R}^2}f(x)f(y)\sigma\mathrm{e}^{-\frac{\sigma^2}{2}(x-y)^2}\mathrm{d}x\mathrm{d}y\\ &=\int_{\mathbf{R}^2}\mathrm{e}^{-z^2/2}f(z/\sigma+y)f(y)\mathrm{d}z\mathrm{d}y,\end{aligned}$$

先设 f 还是连续且有界的，当 $\sigma\longrightarrow\infty$ 时，右边的极限是 $\sqrt{2\pi}\|f\|_{L^2}$, 由 Fatou 引理推出 $|\phi|$ 是平方可积的. 再用控制收敛定理推出结论. 最后用逼近的方法证明结论对一般的密度函数成立.

15. d- 维随机向量 $(\xi_1,\cdots,\xi_d)$ 服从正态分布当且仅当 $\xi_1,\cdots\xi_d$ 的任何线性组合也服从正态分布.
16. (1) 一个服从正态分布的 d- 维随机向量 $(\xi_1,\cdots,\xi_d)$ 独立 (即 $\xi_1,\cdots,\xi_d$ 互相独立) 当且仅当其协方差矩阵是对角型的. (2) 如果 $\overrightarrow{\xi}=(\xi_1,\cdots,\xi_d)$ 服从正态分布，那么存在正交矩阵 Q 使得随机向量 $\overrightarrow{\xi}Q$ 是独立的.
17. 一个随机变量 ξ 是对称的，如果对任何 $x\in\mathbf{R}$, $\mathsf{P}(\xi>x)=\mathsf{P}(\xi<-x)$. 证明： ξ 是对称的当且仅当特征函数是实值的.
18. 设随机序列 ξ_n 依分布收敛于 ξ, 证明：对应的特征函数列在任何有界区间上一致收敛.
19. 设
$$p(x,y)=\frac{1}{4}(1+xy(x^2-y^2))1_{[-1,1]}(x)1_{[-1,1]}(y),$$
(1) 证明： p 是一个二维密度函数. (2) 如果 (X,Y) 的联合密度是 p, 证明：
$$\mathsf{E}e^{\mathrm{i}(X+Y)\xi}=\mathsf{E}\mathrm{e}^{\mathrm{i}X\xi}\cdot\mathsf{E}\mathrm{e}^{\mathrm{i}Y\xi},\ \xi\in\mathbb{R}.$$
20. 设 ξ 服从参数为 t 的 Cauchy 分布.
    (a) 利用留数定理和控制收敛定理直接计算 ξ 的特征函数;
    (b) 证明：两个独立的服从参数分别为 t,s 的 Cauchy 分布的随机变量之和是一个服从参数为 $t+s$ 的 Cauchy 分布.
    (c) 验证如果 ξ 服从参数 1 的 Cauchy 分布，则 2ξ 的特征函数是 ξ 的特征函数的平方, 以此说明两非独立随机变量和的特征函数可以是两者的乘积.
21. 是否存在独立同分布随机变量 X,Y 使得 $X-Y$ 是 $[-1,1]$ 上均匀分布？
    ►不存在. 如果存在的话，那么对任何 $x\in\mathbf{R}$,
$$|\phi(x)|^2=\int_{-1}^{1}\mathrm{e}^{ixy}\frac{1}{2}\mathrm{d}y=\frac{\sin x}{x}$$
    这显然是不可能的.
22. 设 F 是随机变量 ξ 的分布函数. 证明：下面两个等式
$$\lim_{T\longrightarrow\infty}\frac{1}{2T}\int_{-T}^{T}\mathrm{e}^{-iax}\hat{F}(x)\mathrm{d}x=\mathsf{P}(\xi=a);$$
$$\lim_{T\longrightarrow\infty}\frac{1}{2T}\int_{-T}^{T}|\hat{F}(x)|^2\mathrm{d}x=\sum_{x\in\mathbf{R}}(\mathsf{P}(\xi=x))^2.$$
23. 设 ϕ 是随机变量 ξ 的特征函数，如果 $\phi''(0)=0$, 证明： ξ 必定是常数. 由此证明当 $n>2$ 时函数 $x\mapsto\mathrm{e}^{-|x|^n}$ 不是一个特征函数.

▶设 ϕ 是 ξ 的特征函数，如果 ϕ 在 0 点二次可导，则 ξ 平方可积且 $\mathsf{E}\xi^2 = -\phi''(0)$. 事实上，由 Fatou 引理

$$\begin{aligned}
-\phi''(0) &= \lim_{h\to 0}\frac{2\phi(0)-\phi(h)-\phi(-h)}{h^2} \\
&= \lim_{h\to 0}\frac{1}{h^2}\int 2(1-\cos xh)\mathrm{d}F(x) \\
&\geq 2\int \liminf \frac{1-\cos xh}{h^2}\mathrm{d}F(x) = \int x^2\mathrm{d}F(x) = \mathsf{E}\xi^2,
\end{aligned}$$

因此 ξ 平方可积. 如果 $\phi(x) = \mathrm{e}^{-|x|^n}$ 是特征函数，那么 $\phi''(0) = 0$, 因此得 $\mathsf{E}\xi^2 = 0$, 即 $\xi = 0$, 矛盾.

24. 设 F 是分布函数. 证明：如果 $\int |\hat{F}(x)|^2\mathrm{d}x < \infty$, 则 F 连续.

▶利用恒等式 $\int_{\mathbf{R}}\frac{1-\cos ax}{x^2}\mathrm{d}x = a\pi$.

$$\begin{aligned}
\left|\int_{-T}^{T}\frac{\mathrm{e}^{-iax}-\mathrm{e}^{-ibx}}{ix}\hat{F}(x)\mathrm{d}x\right|^2 &\leq \int_{-T}^{T}|\hat{F}(x)|^2\mathrm{d}x\cdot\int_{-T}^{T}\left|\frac{1-\mathrm{e}^{i(b-a)x}}{x}\right|^2\mathrm{d}x \\
&\leq \int_{\mathbf{R}}|\hat{F}(x)|^2\mathrm{d}x\cdot\int_{\mathbf{R}}\frac{2(1-\cos(b-a)x)}{x^2}\mathrm{d}x \\
&\leq C\cdot(b-a),
\end{aligned}$$

由此推出连续性.

▶另外还有一个公式：设 F 是 ξ 的分布函数

$$\begin{aligned}
\int_{-T}^{T}\mathrm{e}^{-iax}\hat{F}(x)\mathrm{d}x &= \mathsf{E}\int_{-T}^{T}\mathrm{e}^{i(\xi-a)x}\mathrm{d}x \\
&= \mathsf{E}\left(\int_{-T}^{T}\mathrm{e}^{i(\xi-a)x}\mathrm{d}x;\xi\neq a\right) + 2T\mathsf{P}(\xi=a) \\
&= \mathsf{E}\left(\frac{\sin(\xi-a)T}{\xi-a};\xi\neq a\right) + 2T\mathsf{P}(\xi=a),
\end{aligned}$$

由控制收敛定理推出公式

$$\lim_{T\longrightarrow\infty}\frac{1}{2T}\int_{-T}^{T}\mathrm{e}^{-iax}\hat{F}(x)\mathrm{d}x = \mathsf{P}(\xi=a).$$

然后设 ξ,η 独立有同样的分布 F, 对 $\xi-\eta$ 应用上面的公式， $a=0$,

$$\lim_{T\longrightarrow\infty}\frac{1}{2T}\int_{-T}^{T}|\hat{F}(x)|^2\mathrm{d}x = \mathsf{P}(\xi=\eta) = \sum_{x\in\mathbf{R}}(\mathsf{P}(\xi=x))^2.$$

由此立刻看出习题的结论成立.

25. 设 X, Y 独立服从标准正态分布，证明： $X+Y$ 与 $X-Y$ 独立.
26. 设有独立同分布的平方可积且标准化的随机变量 X, Y, 如果 $X+Y$ 与 $X-Y$ 独立，证明： X, Y 服从标准正态分布. (提示：设 X, Y 的特征函数是 ϕ. 首先证明 ϕ 恒不为零. 再证明 ϕ 是实的，这由考虑商 $p(x)=\phi(-x)/\phi(x)$ 并由性质 $p(2x)=p(x)^2$ 推出.)

   ▶因为 $(X+Y)+(X-Y)=2X$, 故 $\phi(2x)=\phi(x)^2\cdot\phi(x)\cdot\phi(-x)=\phi(x)^2\cdot|\phi(x)|^2$.

   (1) ϕ 不会等于 0. 如果 $\phi(a)=0$, 那么推出 $\phi(a/2)=0$, 故 $\phi(a/2^n)=0$ 推出 $\phi(0)=0$ 矛盾.

   (2) 令 $p(x)=\phi(-x)/\phi(x)$, 那么 $p(2x)=p(x)^2$. 故 $p(x)=(p(x/2^n))^{2^n}$. 而 $\phi(x)=1-\frac{1}{2}x^2+o(x^2)$, 那么 $p(x)=1+o(x^2)$. 因此 $p(x)=\lim_n(p(x/2^n))^{2^n}=1$, 即 ϕ 是实的.

   (3) 现在 $\phi(x)=\phi(x/2)^4=\phi(x/2^n)^{4^n}\longrightarrow\exp\{-\frac{1}{2}x^2\}$.
27. 设 X 服从 $\Gamma(m,\alpha)$ 分布， Y 独立于 X 服从参数为 $n, m-n$ 的 β 分布， m, n 是非负整数， $n\le m$. 用特征函数的方法求 XY 的分布.
28. 设 X_n 是 $\{1,2,\cdots,n\}$ 上均匀分布的，证明：对任何 $y\in[0,1]$, $\lim_n \mathsf{P}(X_n\le ny)=y$.

   ▶证明 X_n/n 弱收敛于 $[0,1]$ 上均匀分布.

$$\mathsf{E}\mathrm{e}^{ixX_n/n}=\sum_{k=1}^{n}\frac{1}{n}\mathrm{e}^{ixk/n}=\frac{1}{n}\frac{(\mathrm{e}^{ix}-1)\mathrm{e}^{ix/n}}{\mathrm{e}^{ix/n}-1}\longrightarrow\frac{\mathrm{e}^{ix}-1}{ix},$$

   右边是 $[0,1]$ 上均匀分布的特征函数.
29. 设随机变量 X_n 的分布函数为

$$F_n(x)=x-\frac{\sin(2n\pi x)}{2n\pi},\quad x\in[0,1].$$

   (1) 证明： F_n 是分布函数且它有密度函数. (2) 证明：当 $n\to\infty$ 时， F_n 收敛于均匀分布函数，而其密度函数不收敛于均匀分布的密度函数.
30. 用中心极限定理证明

$$\lim_n \mathrm{e}^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}=\frac{1}{2}.$$

   ▶设 ξ_n 独立服从参数 1 的 Poisson 分布. 那么中心极限定理

$$\mathsf{P}(\frac{1}{\sqrt{n}}(\sum_{k=1}^{n}\xi_k-n)\le 0)\longrightarrow\Phi(0)=\frac{1}{2},$$

而 $\sum_{k=1}^{n}\xi_k$ 服从参数为 n 的 Poisson 分布，左边等于

$$\mathsf{P}(\sum_{k=1}^{n}\xi_k \le n) = \mathrm{e}^{-n}\sum_{k=1}^{n}\frac{n^k}{k!}.$$

31. 证明：如果 $\hat{F}_n$ 点点收敛，那么 F_n 弱收敛.
32. 用特征函数方法证明 Khinchin 大数定律：可积的独立同分布随机序列 $\{\xi_n\}$ 服从大数定律.